AD_____

Award Number: W81XWH-04-1-0078

TITLE: Periscopic Spine Surgery

PRINCIPAL INVESTIGATOR: Kevin R. Cleary, Ph.D.

CONTRACTING ORGANIZATION: Georgetown University
Washington DC 20057

REPORT DATE: January 2007

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 01-01-2007	2. REPORT TYPE Annual	3. DATES COVERED *(From - To)* 22 Dec 05 - 21 Dec 06

4. TITLE AND SUBTITLE
Periscopic Spine Surgery

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-04-1-0078

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kevin R. Cleary, Ph.D.

E-Mail: cleary@georgetown.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Georgetown University
Washington DC 20057

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Periscopic Spine Surgery project was envisioned to lay the ground work for developing the physician assist systems of the future. These systems will incorporate robotics, tracking, and visualization to improve the precision of instrument placement and manipulation in minimally invasive procedures. This project has enabled the Georgetown team to become a world leader in the emerging fields of computer aided surgery and medical robotics. Our goal will continue to be to develop systems to add the physician in these demanding minimally invasive procedures with the ultimate aim of improving patient care. Key research accomplishments included: 1. Developed a novel end-effector and automated registration technique for instrument guidance during minimally invasive procedures using a portable needle driver robot. 2. Developed a prototype exoskeleton for shoulder therapy during rehabilitation. 3. Completed a clinical trial on sensorimotor adaptation using a forearm-based robotics system. 4. Developed an open source software architecture and implementation for rapid prototyping of image-guided surgery systems. 5. Completed an image-guided system for more accurate transbronchial biopsy based on electromagnetic tracking of instruments.

15. SUBJECT TERMS
Minimally invasive procedures, technology development, medical robotics, electromagnetic tracking, workflow, open source software

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	117	19b. TELEPHONE NUMBER *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

1   Table of Contents

List of Figures

2   Introduction

This project is aimed at improving the state of the art of image-guided and minimally invasive procedures by developing a new generation of clinical techniques along with the computer-based hardware and software needed for their implementation. The focus of the project is to develop physician assist systems incorporating robotics, tracking, and visualization to improve the precision of instrument placement and manipulation in minimally invasive procedures. This goal is accomplished through a technology development process starting with phantom studies, then proceeding to animal studies, followed by the ultimate aim of human clinical trials.

The project is led by the Imaging Sciences and Information Systems (ISIS) Center of the Department of Radiology at Georgetown University. Project collaborators include the Department of Radiation Medicine at Georgetown University Hospital, the Urology Robotics Group at Johns Hopkins Medical Institutions, the NSF sponsored Engineering Research Center for Computer Integrated Surgical Systems and Technology at Johns Hopkins University, and the Engineering School at the Catholic University of America, as well as several commercial collaborators. Commercial collaborators include Kitware Inc., Traxtal Inc., Accuray Inc., and Siemens Medical Systems. The funds provided under the research are leveraged with funds from other research supporters to create a synergistic environment for advancement of these techniques.

3   Report Body

This section describes the research accomplishments associated with each of the tasks in the statement of work. A brief overview is given and the reader will be referred to the published papers in the appendix for more details. This is an annual report and includes research performed from December 21, 2005 to December 20, 2006. The award number is W81XWH-04-1-0078.

3.1 Task 1: Needle Driver Robotics

Our research group has been a world leader in the development of medical robotics for the interventional environment. Under previous Army funding, we conducted the world's first clinical trial of using a needle driver robot for spinal nerve blocks (Figure 1 and Figure 2). This was a collaborative effort with the URobotics laboratory at Johns Hopkins Medical Institution.
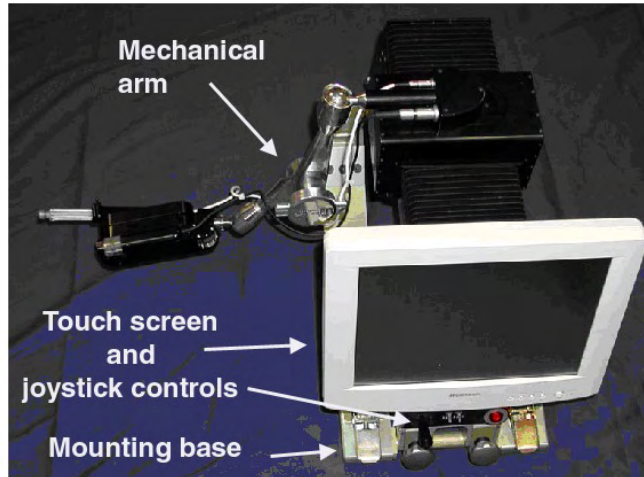
Figure 1: Robotic device showing mechanical arm and joystick control
(courtesy of Dan Stoianovici, PhD, Johns Hopkins Urology Robotics)

**Figure 2: Clinical trial of robotic device for nerve and facet blocks at Georgetown
University (interventional radiologist is Vance Watson, MD)**

In addition, the principal investigator Dr. Cleary has taken a leadership role in the field
and organized several workshops at leading conferences. Most recently, he organized a
Medical Robotics Workshop at the October 2007 MICCAI meeting in Copenhagen.

To continue pushing the boundaries of this field, we have been working with two partners
that have developed different robotic systems for interventional procedures: Johns
Hopkins in Baltimore and Seisdorfer Research in Austria. We have been focusing on two
subtasks in this current effort:
1. integrating a new end-effector from Johns Hopkins into our robotic lung biopsy
    testbed
2. extending a small needle driver robot from the Austrian group for precision
    placement under CT fluoroscopy

Subtask 1: Based on our initial work on nerve and facet blocks as shown in Figure 2, we expanded the application of the Hopkins robot to lung biopsy under CT fluoroscopy. Lung biopsy is an important clinical problem as improved lung cancer screening continues to uncover more suspicious nodules that need to be biopsied. Successful biopsies can be difficult is the nodules are small. Therefore, the development of a robotic system that improves the biopsy of small lesions has high clinical significance.

The goal of this subtask is to integrate a new end-effector into the Johns Hopkins needle driver robot and validate the performance of this end-effector in the CT environment [Stenzel 2006]. The performance will be validated using the lung biopsy software developed at the ISIS Center and shown in Figure 3. An abdominal phantom with precision target points will be used to provide ground truth for the experiments.
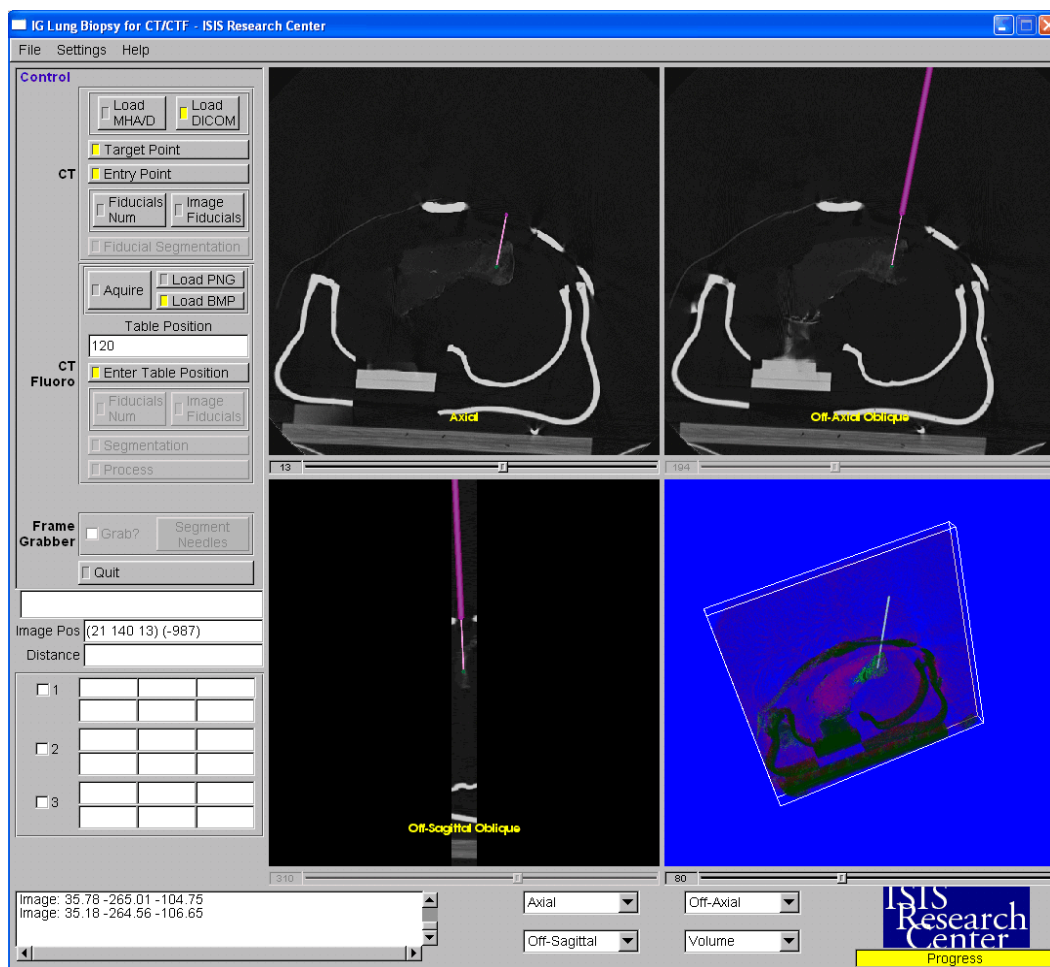
Figure 3: Biopsy software for targeting lesions using CT fluoroscopy

The experimental protocol is as follows:
- Place the robot and phantom on the CT table
- Scan the phantom using 1 mm axial slices
- Transfer the CT scan to a planning workstation running the lung biopsy software shown in Figure 3

- The physician will then select the skin entry point and target point
- The robot will be commanded to spin the needle and drive the needle to the target point (we anticipate the spinning end-effector will reduce friction and enable better straight-line trajectories)
- Another CT scan will be taken to determine the accuracy of the final needle position

The new end-effector has just been fabricated by our colleagues at Johns Hopkins and they are updating the control hardware and software to support this end-effector. We anticipate this work will be complete by Spring 2007 and testing will start immediately thereafter.

Subtask 2: The goal of this subtask is to extend the capabilities of a modular needle driver robot from Seisdorfer Research in Austria to provide automatic registration under CT fluoroscopy. The end-effector of this robot incorporates a spiral fiducial pattern for registration as shown in Figure 4. The software to automatically register the robot has been completed and an evaluation study was carried out as described in [Stenzel 2007]. To evaluate the algorithmic accuracy, we obtained CT scans from different robot positions and calculated the error based only on these CT scans. The mean error for the overall accuracy evaluation was 0.87 mm for translation and 1.46 degrees for rotation. We also estimated the error between target and instrument tip in a worst case scenario as 3.42 mm at an instrument insertion depth of 10 cm.

Figure 4: Spiral fiducial pattern on robot end-effector for auto registration

3.2 Task 2: Rehabilitation Robotics

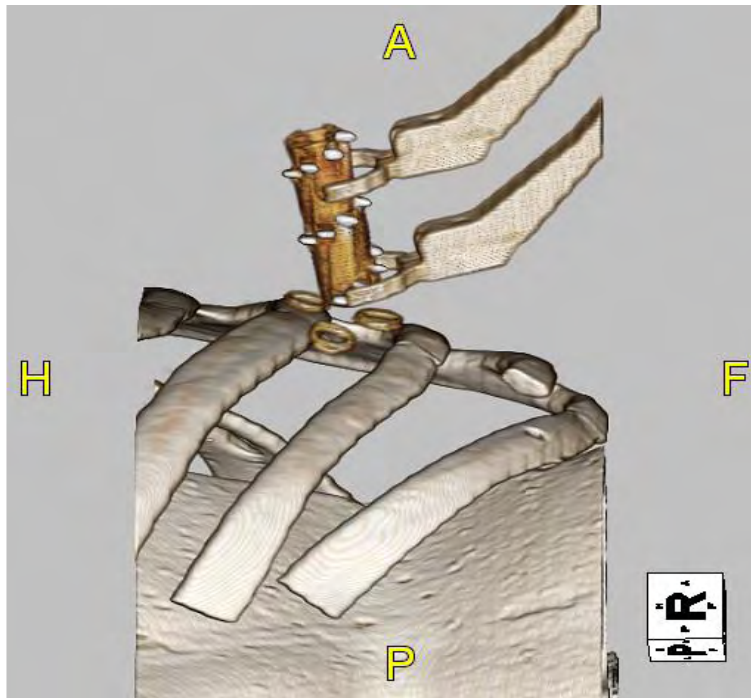The primary goal of this effort is to develop robotic systems for physical therapy and neurorehabilitation. The secondary goal is to expand this capability to the Internet. We currently have two rehabilitation robotic testbeds: the Maryland-Georgetown-Army (MGA) Exoskeleton and a pair of InMotion2 shoulder-elbow robots. This has been a collaborative effort with the National Rehabilitation Hospital, the University of Maryland, the Center of Excellence for Remote and Medically Under-Served Areas (CERMUSA), and Interactive Motion Technologies, Inc. We focused on two subtasks here:
1. Shoulder therapy with the MGA exoskeleton
2. Neurorehabilitation using InMotion2 robot

Subtask 1: Shoulder therapy. The shoulder is the most mobile joint in the body and consequently, has the highest incidence of injury in the upper extremity. The purpose of this project is to develop protocols for a robotic arm exoskeleton that will facilitate precise multimode control of patient arm movements to tailor rehabilitation programs for optimum effect. Exoskeletons can modulate the level of assistance over multiple directions thereby allowing varying degrees of resistance while performing the task. This capability offers a significant advantage over traditional linear and cam exercise devices that restrict movement to a single direction.

We have built a robotic arm exoskeleton for shoulder rehabilitation, which is currently undergoing integration and testing (Figure 5) [Carignan 2007]. The mechanical hardware was designed and built by the University of Maryland Space Systems Laboratory, and CAIMR is developing the electronics, control system, and therapy protocols. Although the exoskeleton has achieved basic operation (Figure 6), a number of redundant sensors need to be installed to fulfill the safety requirements. We expect the system to become fully operational by the end of 2007.

Subtask 2: Neurorehabilitation. We have just completed a clinical trial on sensorimotor adaptation when subjects are exposed to kinematic and force distortions applied by the IM2 Robot (Figure 7). By examining the response of both healthy and afflicted subjects to novel kinematic and/or dynamic environments, a quantitative empirical portrait of neural algorithms can be formulated. The parameters within these internal models can then be fine-tuned to study how they produce a given neurological condition. This has potential to broadly impact the rehabilitation and treatment of patients afflicted with neurological conditions.

The 30 participants in this study were all healthy, adult volunteers without any known neurological deficits. Although this trial involved testing healthy subjects, the eventual goal of this project is to develop therapy for subjects with Parkinson's Disease. Preliminary results from both the kinematic and force distortion trials show strong evidence of subject learning and adaptation (see Figure 8). Results from analyses on the data from these trials and have submitted to a journal paper [Tang 2007].
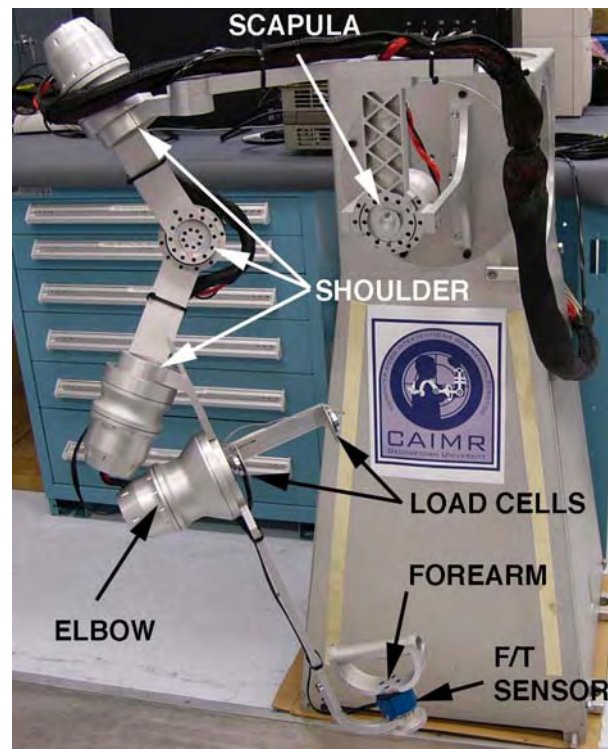
Figure 5: Exoskeleton hardware showing location of sensors and actuators

Figure 6: Exoskeleton worn by subject during a fit check
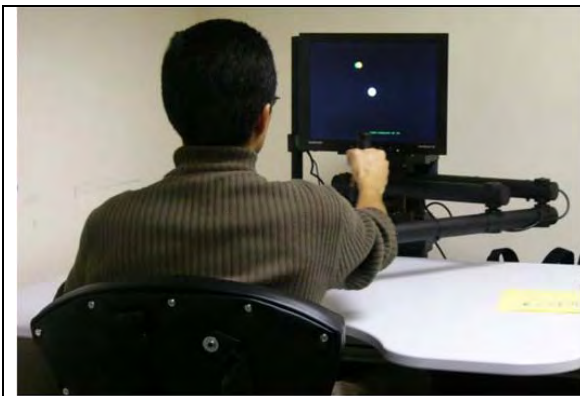
Figure 7: Inmotion2 Robot operator interface being used for sensorimotor clinical trial	Figure 8: Results from kinematic distortion trials show evidence of learning
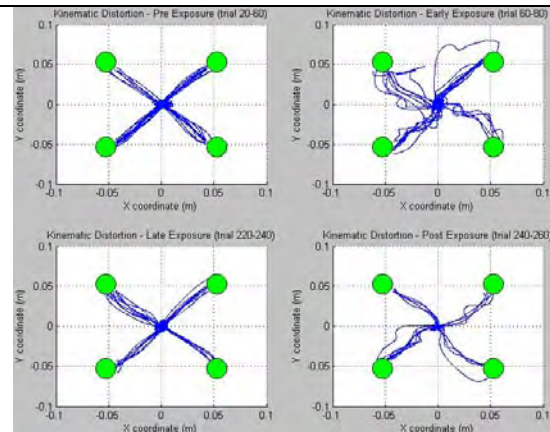
3.3 Task 3: Interventional Suite of the Future

This task is an outgrowth of the OR2020 workshop / Operating Room of the Future meeting that was organized by Dr. Cleary in March 2004 and co-sponsored by TATRC/USAMRMC. At the meeting, one of the identified needs was for standards for integrating hardware and software in the operating room and interventional suite for better patient care. The goal of the Interventional Suite of the Future project at Georgetown is to develop a testbed to help bring these concepts to fruition.

Supported in part by the Periscopic Spine Surgery project, the Department of Radiology at Georgetown University Hospital just finished installing a state-of-the-art Interventional Suite. This suite consists of a C-arm angiography system with a flat panel detector, which provides excellent soft tissue contrast for interventional procedures. In the past year, we focused on two subtasks here:

1. Integrate electromagnetic tracking for image guidance of abdominal procedures based on open source software
2. Develop PET/CT image fusion for image-guided biopsy, treatment, and therapy evaluation of cancer

Subtask 1: The goal of this subtask was to integrate the image-guided system we have developed over the past five years into the new Interventional Suite. The image-guided system includes an electromagnetic tracking system and electromagnetically tracked needles to provide a virtual image overlay and guidance capability. This work builds on a large body of preliminary work, namely:

a) A study to develop a standard hardware and software platform for the evaluation of electromagnetic tracking accuracy in the clinical environment [Wilson 2007]
b) The development of a prototype system for treatment planning and image guidance for radiofrequency ablation of liver tumors [Zhang 2007]

c) The development of a prototype system for integrating ultrasound imaging and image overlay using electromagnetic tracking [Zhang 2006a]

The software has been developed over the past two years as part of the Image-guided Surgical Toolkit (IGSTK) project, an open source effort lead by Dr. Cleary with joint Army and NIH funding (Figure 11). The software has been integrated into the new Interventional Suite as follows:

- A model 57 abdominal phantom (CIRS Inc.) was used in the evaluation
- A rotational angiography image acquisition is obtained followed by a DynaCT reconstruction to obtain CT-like DICOM images
- The images are sent from the aniography system to our computer workstation
- Registration is performed by identifying fiducial points in electromagnetic tracking space and CT space
- Image overlay and instrument tracking are enabled
- The physician can verify the accuracy of the system by targeting known landmarks using image guidance alone and checking the accuracy of final needle placement
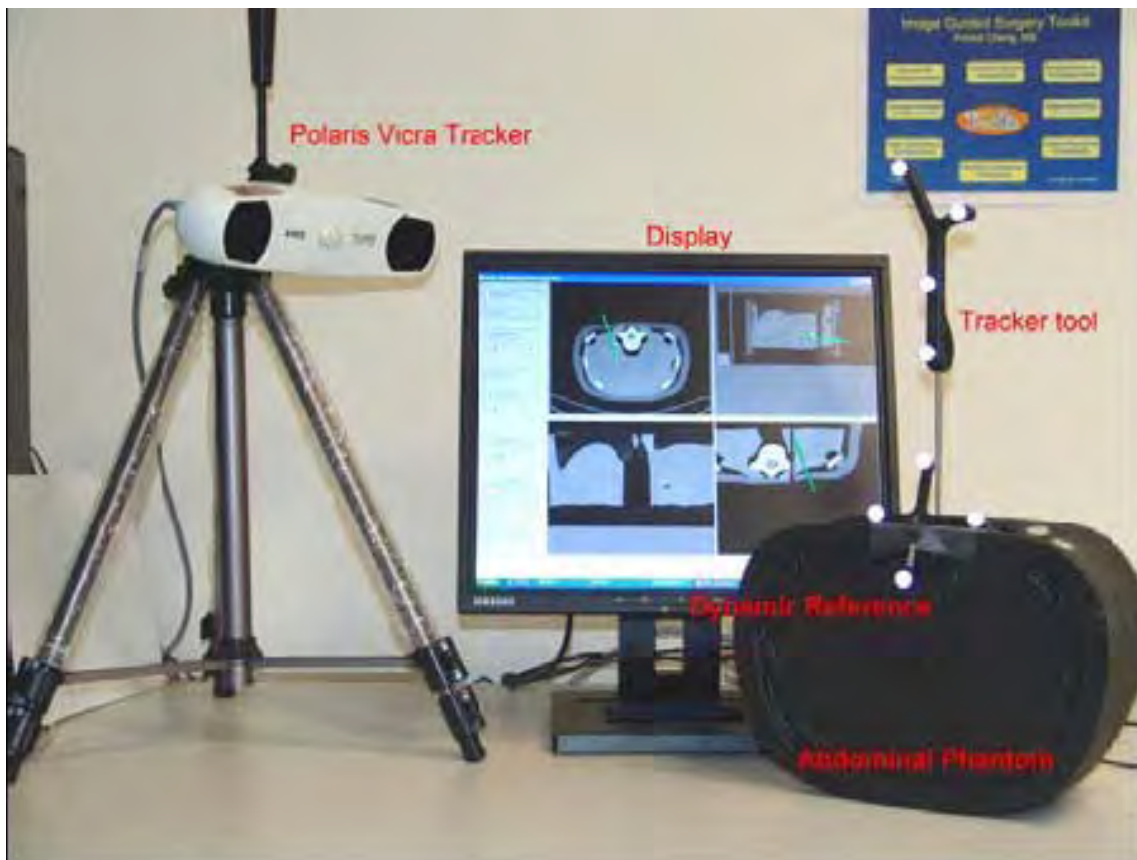

Figure 9: Image-guided software and needle tracking in phantom study

Subtask 2: Minimally invasive, image-guided interventions hold much promise for biomedical science, offering advantages such as more precise tissue sampling and treatment of disease at early stages with high success rates and low morbidity. However, these interventions can only perform as well as the information or images that are used to direct them. The vast majority of interventional radiology procedures are performed under fluoroscopy or CT fluoroscopy, which provides only structural landmarks in the body because of the inherent limitations of x-ray imaging. Providing interventional radiologists with functional information such as metabolic activity, gene expression, or receptor density levels would be a significant and enabling technology because it would allow physicians to directly target areas of biochemical abnormality. Although PET (positron emission tomography) has been widely adopted as a functional imaging modality, it has not yet been effectively integrated into the interventional radiology environment.

The goal of this subtask is therefore to develop a capability for fusing pre-procedure PET/CT images to images obtained in the interventional radiology environment, to enable the use of PET information in targeting lesions. Our initial work here is as follows:

    a) We completed a study to evaluate the influence of CT based attenuation correction on PET/CT registration [Yaniv 2007]

    b) We developed a method to create 4D imaging data using open source image registration software [Wong 2006]

The current PET/CT scanner at Georgetown University Hospital (Siemens Biograph 6) does not have the capability to perform respiratory gating, but this capability is available as a field upgrade. We are awaiting a quotation from the vendor to upgrade this system. With this equipment, it will be possible to generate a series of PET images that represent the extent of the respiratory cycle, from end-inhale to end-exhale.

To test our algorithms, we will use a respiratory motion phantom that can be programmed to emulate a wide variety of respiratory patterns. The phantom is shown in Figure 9. We have already used this phantom for experiments in the nuclear medicine clinic and have acquired PET/CT images from it.

Figure 10: Respiratory motion phantom being scanned in Siemens Biograph at Georgetown University Medical Center during initial PET/CT experiments (phantom constructed by Dan Stoianovici at Johns Hopkins under previous Army funding)

3.4 Task 4: Image-Guided Bronchoscopy

This task is an outgrowth of our work in electromagnetic tracking for transbronchial needle aspiration (TBNA). This is in collaboration with Eric Anderson, MD, a pulmonologist at Georgetown University Hospital. TBNA is a common method to evaluate mediastinal and hilar lymphadenopathy, especially for staging of lung cancer. The procedure has technical limitations, mostly related to the difficulty of accurately placing the biopsy needles into the target lesion. Currently, pulmonologists plan TBNA by examining a number of Computed Tomography (CT) scan slices before the procedure. Then, they manipulate the bronchoscope down the respiratory track and perform biopsies based upon this pre-planning stage, but they have no on-line image guidance during the biopsy. The diagnostic yield of conventional TBNA has been reported to be up to 71% and we believe image guidance may be able to increase this yield.

The goal of this task is to build and test an image-guided system for transbronchial needle aspiration. The system will provide an image overlay based on pre-procedure CT images. The system concept is shown in Figure 11. The system developed to date works as follows:

1) Obtain a pre-operative CT scan of the bronchial area
2) The pulmonologist manually segments the lymph nodes
3) Path planning to the biopsy target is done by the pulmonologist using the reformatted CT images and virtual bronchoscopy images
4) Image registration between CT and patient space is performed using techniques developed in our laboratory
5) The electromagnetic tracking system is enabled to track the tip of the bronoschope as it is advanced through the bronchial tree
6) When the bronchoscope is advanced, three types of images are displayed
    a. corresponding reformatted CT images
    b. transparent virtual bronchoscope images
    c. 3D volume rendered image
7) Using the image guidance described in Step 6, the pulmonologist advances the bronchoscope to the biopsy target and takes the tissue sample

We have completed phantom studies to date and have begun the process of applying for human clinical trials. We have also had two recent related publications:

a) High dynamic range virtual bronchsopy rendering for video tracking [Popa 2007]
b) High quality GPU rendering with displaced pixel shading [Zhang 2006b]
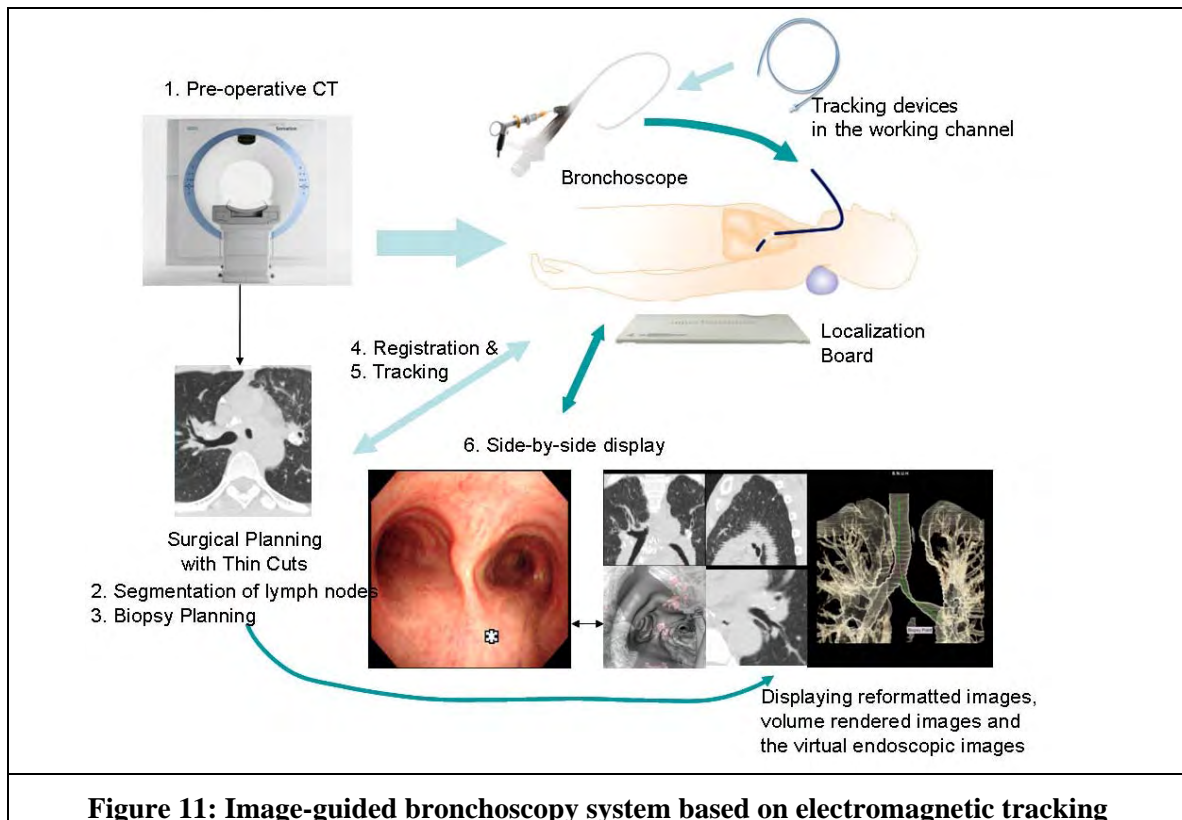
Figure 11: Image-guided bronchoscopy system based on electromagnetic tracking

4   Key Research Outcomes

This section provides a bulleted list of key research accomplishments during the entire project:

1. Completed developing and testing a software environment for targeting of lesions using CT fluoroscopy (Figure 3)

2. Developed a novel end-effector and automated registration technique for instrument guidance during minimally invasive procedures using a portable needle driver robot (Figure 4)

3. Developed a prototype exoskeleton for shoulder therapy during rehabilitation (Figures 5 and 6)

4. Completed a clinical trial on sensorimotor adaptation using a forearm-based robotics system (Figures 7 and 8)

5. Developed an open source architecture and implementation for rapid prototyping of image-guided surgery systems (Figure 9)

6. Began initial study of PET/CT respiration compensation using a novel anthropomorphic phantom developed under previous work (Figure 10)

7. Completed an image-guided system for more accurate transbronchial biopsy based on electromagnetic tracking of instruments (Figure 11)

5   Reportable Outcomes

This section provides a list of reportable outcomes. The major product of this year is the list of papers given in Section 7 (References). Copies of these documents are provided in the appendix.

In addition, several proposals to the National Institutes of Health were submitted based on this work. Based on our open source work of the past several years, we will soon be awarded an NIH R01 proposal to continue this project.

A visiting researcher from Romania was supported under this project. Lucian Grunionu, PhD, is a professor of engineering at the University of Craiova. He spent several months in our lab in the fall of 2006 developing tracked instruments for the image-guided bronchoscopy project.

The research group at Georgetown continued to take a lead in the Washington Area Computer Aided Surgery Society (www.washcas.org), which was formed in 2006 to promote research in this field.

6 Conclusions

The Periscopic Spine Surgery project has continued to lay the ground work for developing the physician assist systems of the future. These systems will incorporate robotics, tracking, and visualization to improve the precision of instrument placement and manipulation in minimally invasive procedures.

We continued our work in medical robotics and demonstrated the ability of these systems to precisely orient instruments for medical procedures. We also continued our work in electromagnetic tracking of instruments and recently applied for FDA approval of a clinical trial for lung biopsy. We have also continued our work on respiratory motion compensation in investigating respiratory motion tracking for the interventional suite, based on rotational angiography and tomographic reconstruction of images.

This project has enabled the Georgetown team to become a world leader in the emerging fields of computer aided surgery and medical robotics. Our goal will continue to be to develop systems to aid the physician in these demanding minimally invasive procedures with the ultimate aim of improving patient care.

7   References

[Carignan 2007]   Craig Carignan, Jonathan Tang, Stephen Roderick and Michael Naylor, "A Configuration-Space Approach to Controlling a Rehabilitation Arm Exoskeleton," to be published in *Proc. IEEE Int. Conf. on Rehabilitation Robotics (ICORR),* January 2007.

[Popa 2007]   Teo Popa and Jae Choi, "High Dynamic Range (HDR) Virtual Bronchoscopy Rendering for Video Tracking," presented at the *SPIE Medical Imaging Symposium* and to be published in the *Visualization and Image-Guided Procedures* track, San Diego, CA, 18-22 February 2007.

[Stenzel 2006]   Roland Stenzel, Filip Banovac and Kevin Cleary, "Augmented Lung Biopsy Under CT Fluoroscopy: Workflow and Registration," poster presentation at *Computer Aided Radiology and Surgery,* Osaka, Japan, 2006.

[Stenzel 2007]   Roland Stenzel, Ralph Lin, Peng Cheng, Gernot Kronreif, Martin Kornfeld, David Lindisch and Kevin Cleary, "Precision Instrument Placement Using a 4-DOF Robot with Integrated Fiducials for Minimally Invasive Interventions," poster presentation to be published in *SPIE Medical Imaging Symposium* within the *Visualization and Image-Guided Procedures* track, San Diego, CA, 18-22 February 2007.

[Tang 2007]   Jonathan Tang, Jose Contreras-Vidal and Craig Carignan, "Comparison of Neurosensorimotor Adaptation Under Kinematic and Dynamic Distortions," to be published in *Proc. IEEE Int. Conf. on Rehabilitation Robotics (ICORR),* January 2007.

[Wilson 2007]   Emmanuel Wilson, Ziv Yaniv, Hui Zhang, Christopher Nafis, Eric Shen, Guy Shechter, Andrew Wiles, Terry Peters, David Lindisch and Kevin Cleary, "A Hardware and Software Protocol of the Evaluation of Electromagnetic Tracker Accuracy in the Clinical Environment: A Multi-Center Study," poster presentation to be published in *SPIE Medical Imaging Symposium* within the *Visualization and Image-Guided Procedures* track, San Diego, CA, 18-22 February 2007.

[Wong 2006]   Kenneth Wong, Luis Ibanez, Teo Popa and Kevin Cleary, "Creation of 4D Imaging Data Using Open Source Image Registration Software," *Proc. of SPIE, Medical Imaging;* Vol. 6141, 194-20, March 2006.

[Yaniv 2007]   Ziv Yaniv, Kenneth Wong, Filip Banovac, Elliot Levy and Kevin Cleary, "The Influence of CT Based Attenuation Correction on PET/CT Registration: An Evaluation Study," presented at the *SPIE Medical Imaging Symposium* and to be published in the *Visualization, Image-Guided Procedures, and Display* track, San Diego, CA, 18-22 February 2007.

[Zhang 2006a]   Hui Zhang, Filip Banovac, Amy White and Kevin Cleary, "3D Freehand Ultrasound Calibration using an Electromagnetically Tracked Needle," *Proc. of SPIE, Medical Imaging;* Vol. 6141, 775-783, March 2006.

[Zhang 2006b]   Hui Zhang and Jae Choi, "High Quality GPU Rendering with Displaced Pixel Shading," *Proc. of SPIE, Medical Imaging;* Vol. 6141, 851-859, March 2006.

[Zhang 2007]   Hui Zhang, Filip Banovac, Stella Munuo, Enrique Campos-Nanez, Hernan Abeledo and Kevin Cleary, "Treatment Planning and Image Guidance for Radiofrequency Ablation of Liver Tumors," poster presentation to be published in *SPIE Medical Imaging Symposium* within the *Visualization and Image-Guided Procedures* track, San Diego, CA, 18-22 February 2007.

8 Appendices (Papers)

Copies of the eleven papers selected for this report are reproduced in this section.

8.1 Carignan 2007: A Configuration-Space Approach …

Reprint begins on the next page and is eight pages.

A Configuration-Space Approach to Controlling a Rehabilitation Arm Exoskeleton

Craig Carignan, Jonathan Tang
Imaging Science and Information Systems (ISIS) Center
Department of Radiology
Georgetown University
Washington, DC 20057
Email: crc32@georgetown.edu, jt96@georgetown.edu

Stephen Roderick, Michael Naylor
Space Systems Laboratory (SSL)
Department of Aerospace Engineering
University of Maryland
College Park, MD 20742
Email: roderick@ssl.umd.edu, naylor@ssl.umd.edu

Abstract— **A modular approach is proposed for controlling a six degree-of-freedom arm exoskeleton used in shoulder rehabilitation. The degrees of freedom are partitioned into anthropomorphic sets of joints to be controlled which depend on the requirements of the particular rehabilitation prototocol. The controllers can operate in impedance or admittance modes to realize both iso-lateral exercise and functional training protocols. The approach is applied to the MGA Exoskeleton, and some preliminary experimental results are given for operating the arm under admittance control during functional training.**

I. Introduction

Arm exoskeletons developed for upper extremity rehabilitation require different approaches to control than those used for virtual reality applications. This paradigm shift is due to the distribution of command forces imparted along discrete points on the arm rather than just at the handle. This need is even more acute when the exoskeleton is kinematically redundant for the task, therefore requiring the control of additional degrees of freedom not directly involved in meeting the primary objective.

As an example, the MGA Exoskeleton shown in Figure 1 incorporates an extra joint to allow for "tilting" or rotation of the scapula joint on the thorax. This rotation causes the shoulder glenohumeral "ball-and-socket" joint to elevate $10 - 12$ cm (up to $60°$) during shoulder shrug or shoulder abduction [1]. While this humeral motion may not be important for virtual reality applications, scapulo-humeral rhythm is essential for shoulder rehabilitation [2], [3]. Other devices such as the *ESA Human Arm Exoskeleton* have also incorporated shoulder translation in their designs [4].

The additional degrees of freedom (DOF) can render the traditional Cartesian workspace approach inadequate for rehabilitation applications. This inadequacy leads to consideration of alternative "configuration-control" approaches previously used for controlling excess degrees of freedom in kinematically redundant manipulators [5]. However, the application to exoskeletons is quite different in that the human arm is "encased" in the robot arm and must drive the entire configuration.

The article begins with a brief history of arm exoskeleton control in Section II. A short description of some of the shoulder rehabilitation protocols to be implemented is given in Section III. The configuration space kinematics for the exoskeleton are derived in Section IV. The system architecture and modular control approach are described in Section V. Preliminary results for some admittance control experiments for the functional training mode are presented in Section VI. Discussion of the results and plans for future work are outlined in Section VII.
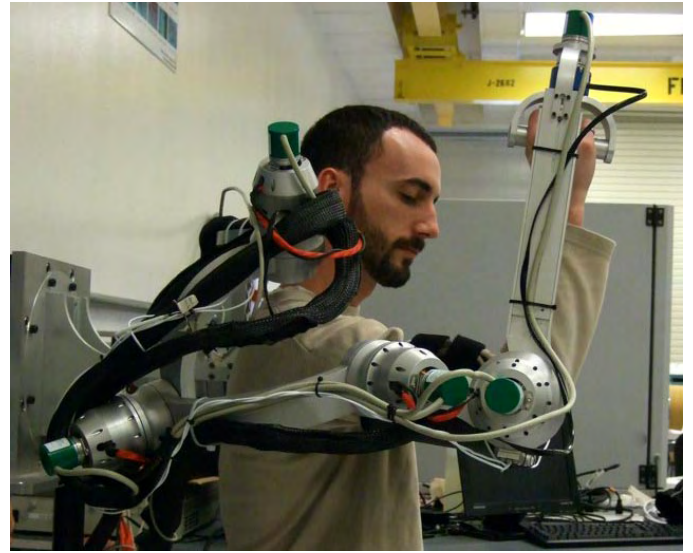
Fig. 1. The MGA Exoskeleton incorporates a scapula tilt axis as well as an orthogonal axis triad for shoulder rotation.

II. Previous Work

Most arm exoskeletons built to-date were developed as either force-reflecting master arms for teleoperation or as haptic devices for virtual reality (VR) applications [6]. The purpose of these exoskeletons is to impart "contact" forces at the handle of the exoskeleton that replicate forces sensed by the slave arm (teleoperation) or interaction with a virtual environment (VR application). Nearly all of these devices implement a basic form of impedance control in which the Cartesian forces at the handle are mapped into joint torque commands using the Jacobian. The *PHANTOM* Haptic Interface is a classic example of a device utilizing this approach [7].

The main advantages of the impedance control approach are: it does not require inverse kinematics, and it works well for low impedances. The main drawback is that good force replication at the handle requires compensation of the natural dynamics of the exoskeleton, such as gravity loading and drive friction. A force loop is sometimes closed around the force sensor to help reduce the impedance errors, but this can destabilize the system [8]. Classic examples which utilize this approach are the *Exoskeleton Arm-Master* [6] and *GLAD-IN-ART Exoskeleton* [9]. A more current example is the *L-Exos* Exoskeleton [10].

An alternative approach called "admittance" control has recently emerged, but mostly in the context of manipulators used as large-reach haptic devices [11]–[13]. In this approach, the sensed force at the exoskeleton gripper is used as the input to a desired impedance model, which outputs a desired motion to be produced at the gripper. This Cartesian position is mapped into joint position commands using the inverse kinematics, which are then input to a proportional-derivative (PD) servocontroller to drive the joints to the desired position.

The admittance control approach has the major advantage of not requiring model feedforward to compensate for the natural dynamics; instead, it relies on the high PD gains of the joint position servo-loop to reject unmodeled dynamics. However, it has the major drawback of instability for high admittance (low impedance), which is the opposite of impedance control [14]. The *Sensor Arm* [15] is a classic example of an exoskeleton implementing this approach. Few arm exoskeletons seem to use admittance control, although the *ARMin Exoskeleton* [16] appears to be able to operate in both admittance and impedance modes.

III. Exoskeleton Exercise Protocols

The MGA Exoskeleton has a total of five actuated joints and a single unpowered joint. An orthogonal, intersecting-axis triad is used to realize glenohumural shoulder rotation as shown in Figure 2. The first shoulder axis joint is mounted at a $30°$ angle to the azimuth axis to rotate the singularity away from the vertical position (alignment of first and third axes). The third shoulder axis intersects the upper arm at an angle of $45°$. A single rotary joint normal to the back approximates shoulder elevation and depression along an arc. A single pitch joint drives elbow flexion/extension, and a passive forearm roll joint aft of the handle allows free forearm supination/pronation.

The range of motion and torque specifications for the exoskeleton versus the average adult male are given in Table I. The range of motion of the exoskeleton is comparable to the human, although direct comparison is difficult due to the serial nature of the exoskeleton shoulder triad compared to the human "ball-and-socket" humeral joint. The torques are comparable to the maximum strength of the average adult male [17].

There are basically two classes of shoulder therapy protocols being implemented on the exoskeleton: iso-lateral exercise and functional training. Iso-lateral exercises are
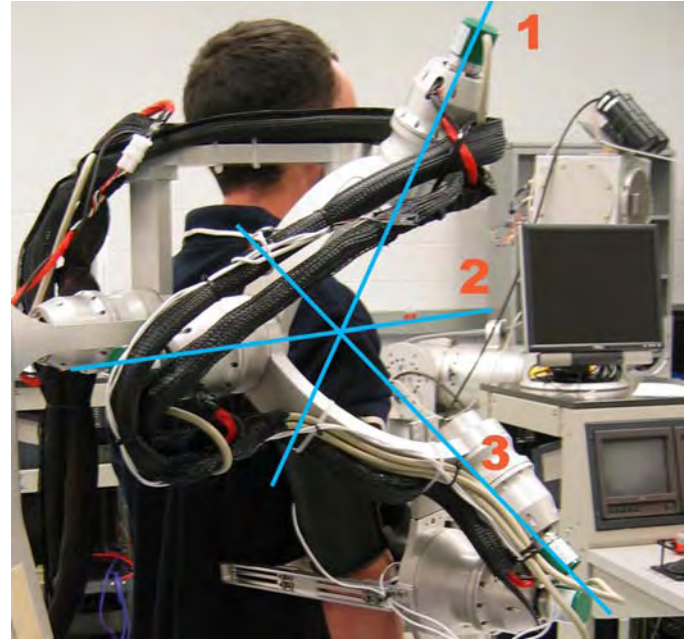

Fig. 2. An intersecting-axis triad is used to approximate three-dimensional rotation of the glenohumeral ball-and-socket joint.

those that occur around a single rotation axis of the shoulder or along a straight line path of the hand. Functional training involves more general movement of the hand through three-dimensional space which occurs in everyday tasks. Each mode is described in more detail below.

TABLE I
RANGE OF MOTION AND TORQUE CAPACITY FOR HUMAN AND MGA EXOSKELETON.

Joint	Joint Range (deg)		Max. Torque (N-m)	
	Human	Exos	Human	Exos
Scapula Elev/Dep	30/30	30/30	?	137
Shoulder Flex/Ext	188/61	147/46	115/110	137
Shoulder Abd/Add	134/48	105/75	134/94	137
Shoulder Med/Lat	97/34	89/63	?	137
Elbow Flex/Ext	142/0	142/0	72/42	69
Forearm Pro/Sup	85/90	180/180	9/7	–

A. Iso-Lateral Exercise

Iso-lateral exercises closely mimic those currently performed manually or with the assistance of exercise machines [2]. Examples of shoulder rotation exercises include internal/external rotation and shoulder abduction/adduction as shown in Figure 3. In rotational exercises, the motion of the shoulder joints is determined by the resistance about the desired shoulder axis of rotation. The scapula joint moves along prescribed path based on biomechanical model of shoulder motion. The elbow pitch is free to move since it does not affect shoulder rotation.

Examples of translational exercises include upright rows and wall push-ups. In these exercises, the hand is restricted to move along a straight line in three dimensions. Therefore, only three out of four shoulder/elbow joints are constrained.

Fig. 3. Shoulder shown at 90° abduction.

B. Functional Training

Functional training has been shown to improve performance of tasks that mirror the same type of movement. The patient views the simulated task and representation of their arm through a head mounted display while the exoskeleton provides haptic feedback to the patient. A force sensor located at the hand gripper senses the forces being exerted by the patient's "contact" with the virtual environment and relays them to the controller which commands the exoskeleton in response to the interaction.

Examples of functional training are proprioceptive neuromuscular facilitation (PNF) patterns and VR task simulations. In PNF tasks, the 3D curvilinear path of the arm is designed to evoke particular activation of muscles intended to simulate tasks [18], [19]. The path of hand and elbow determine the shoulder and elbow angles, and the scapula joint moves along prescribed path based on biomechanical model of shoulder motion. In VR simulations, interaction force at the handle determined by "contact" with virtual world. A compression load cell atop the shoulder senses forces and raises the scapula joint in response.

IV. CONFIGURATION SPACE KINEMATICS

The workspace of the exoskeleton is highly dependent upon the task being performed. In traditional exoskeletons used for VR applications, the goal is to program the interface (handle) so that it feels as though the user is interacting with a virtual environment. In this case, workspace is the three-dimensional position and orientation of the handle. Link frames are assigned as shown in Figure 4 using the Denavit-Hartenburg (D-H) Method [20], and the forward kinematics are performed iteratively from the base to the handle to determine the position and orientation of the handle with respect to the base frame.

The MGA Exoskeleton differs in two significant ways from traditional designs: (a) there is an additional (scapula) joint between the torso and the first shoulder joint, and (b) there are four (not three) degrees of freedom in the arm. These variations introduce questions of how to handle the kinematic redundancy, from both kinematic and practical

standpoints. What should function as the base frame: 0 or B? How should the extra DOF in the arm be specified?

These questions can be answered largely by considering how the exoskeleton will be controlled. Since the scapula joint will be controlled independently from the arm, it made the most sense to keep $\{0\}$ as the base frame. The body frame $\{B\}$ is then defined as a rotation θ_0 about the x_0-axis from $\{0\}$. The link frames for joints 0-5 are then assigned using the D-H Method. Retaining $\{0\}$ as the base frame greatly simplifies the forward kinematics.

The question of how to resolve the excess DOFs is more complex. If the exoskeleton is being used as a VR device, then the Cartesian position of the handle is specified, and the "self-motion" is free. If the exoskeleton is being used for iso-lateral exercise about a particular shoulder axis, then the torques about the three shoulder axes are specified and the elbow pitch is free. The configuration kinematics for both of these cases is examined below.

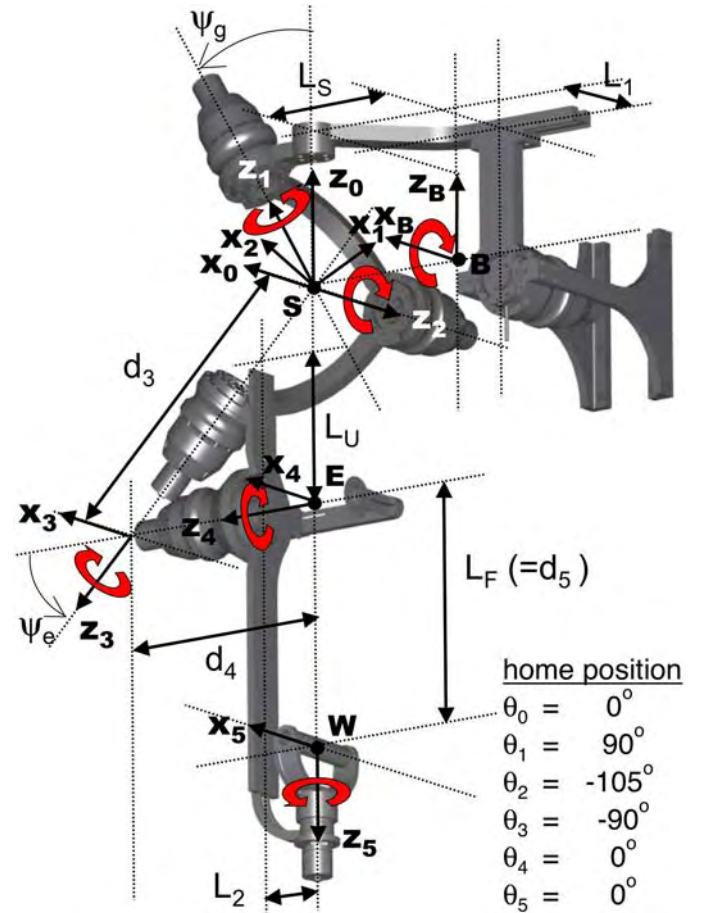Fig. 4. D-H link frame assignment for the MGA Exoskeleton. Axes 1-5 are the arm joints, and axis B is the scapulothoracic joint. Linkages L_F, L_U, and L_S have adjustable lengths.

A. Configuration Space Kinematics

The configuration space for the exoskeleton is comprised of two parts: the Cartesian position of the gripper relative to the body frame, and the "orbit" of the elbow about the

shoulder-wrist axis. The handle position is derived using forward kinematics, and an approach called the "Configuration Control Method" will be used to specify the self-motion of the arm [5].

1) Handle Position: The Cartesian position of the handle relative to the base frame can be found by first assigning D-H link frames as shown in Figure 4. The D-H parameters for the kinematics are given in Table II. [Note that the θ column indicates the value of the joint variable in the home position.] The lengths L designated with a subscript letter are all passive sliding joints, while those with a numbered subscript are fixed: $L_S = 16.35 - 30.00$ cm, $L_E = 19.60 - 23.60$ cm, $L_F = 30.00 - 39.00$ cm, $L_1 = 12.06$ cm, and $L_2 = 7.29$ cm. The angles ψ between the axes are both fixed and have values $\psi_g = 30°$ and $\psi_e = 45°$. Note that value of d_4 is negative since it represents a translation along the negative $\hat{z}_4$-axis.

TABLE II

D-H PARAMETERS FOR THE MGA EXOSKELETON.

link i	a_{i-1} (cm)	α_{i-1} (deg)	d_i (cm)	θ_i* home (deg)
1	0	+30	0	+90
2	0	−90	0	−105
3	0	+90	$\sqrt{2}(L_U + L_2)$	−90
4	0	−45	$-(L_U + L_2)$	0
5	0	−90	L_F	0

Using the D-H parameters, the local link position and rotation transforms, $^i p_{i+1}$ and $^i R_{i+1}$, can be determined for links 1 through 5. These transforms are then cascaded to find the position $^0 p_5$ and orientation $^0 R_5$ of the handle with respect to the base frame [20]. The transforms from the body frame to the base frame are given by:

$$^B p_0 = [0 \; -L_S \; 0]^T \tag{1}$$
$$^B \mathbf{R}_0 = R_X(-\theta_0) \tag{2}$$

where $R_X(\gamma)$ is a pure rotation about the x-axis

$$\mathbf{R}_X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\gamma) & -sin(\gamma) \\ 0 & sin(\gamma) & cos(\gamma) \end{bmatrix} \tag{3}$$

and $\gamma = -\theta_0$. The two sets of transformations can then be cascaded to determine the position of the handle with respect to the body frame

$$^B R_5 = {}^B R_0 {}^0 R_5 \tag{4}$$
$$^B p_5 = {}^B p_0 + {}^B R_0 {}^0 p_5 \tag{5}$$

where $^B p_5$ is the position of the handle.

2) Elbow Orbit Angle: "Self-motion" of a manipulator is defined as the motion that occurs when the end-effector is locked. In the exoskeleton, this "self-motion" is the ability of the elbow to "orbit" about the shoulder-wrist axis p_w shown in Figure 5 while the position of the wrist and shoulder are

held fixed [21]. The orbital angle ϕ is defined as the angle that the plane formed by the points S, E, and W makes with the reference plane defined by the reference vector, $\hat{\mathbf{v}}$, and the shoulder-wrist vector, $\mathbf{p}_w$ [21].
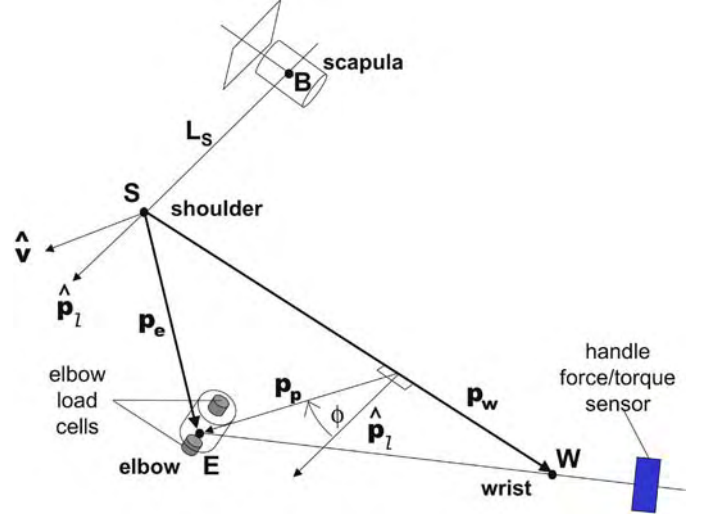

Fig. 5. Definition of elbow orbit angle and location of gripper force/torque sensor and elbow axial load cells.

Let the wrist and elbow vectors be defined as $p_w = {}^0 p_5$ and $p_e = {}^0 p_4$, respectively, and let $\hat{\mathbf{v}}$ denote an arbitrary fixed unit reference vector. Now let the following auxiliary vectors be defined

$$\mathbf{p}_d \equiv \hat{\mathbf{p}}_w (\hat{\mathbf{p}}_w^T \mathbf{p}_e) \tag{6}$$
$$\mathbf{p}_p \equiv \mathbf{p}_e - \mathbf{p}_d \tag{7}$$
$$\mathbf{p}_\ell \equiv (\mathbf{p}_w \times \hat{\mathbf{v}}) \times \mathbf{p}_w \tag{8}$$

The roll angle of the SEW plane or "elbow orbit angle" is defined as the angle between $\mathbf{p}_p$ and $\mathbf{p}_\ell$

$$tan\phi \equiv \frac{\hat{\mathbf{p}}_w^T (\mathbf{p}_\ell \times \mathbf{p}_p)}{\mathbf{p}_\ell^T \mathbf{p}_p} \tag{9}$$

Since the elbow orbit angle is a rather complex function of the arm angles, the utility of a symbolic representation is somewhat limited, and ϕ is more easily computed by performing the numerical vector operation in (9) and then taking the arctangent of the result.

B. Shoulder Space Kinematics

The shoulder space kinematics are defined as the orientation of the upper arm (line segment SE) relative to the body frame $\{B\}$. This orientation can be obtained by rotating frame $\{3\}$ by $45°$ about the x_3-axis as follows:

$$^B \mathbf{R}_U = {}^B \mathbf{R}_0 {}^0 \mathbf{R}_3 R_X(45°) \tag{10}$$

where $^B \mathbf{R}_0$ is given by (2) and $R_X(\gamma)$ by (3) where $\gamma = 45°$. The direction of the upper arm in the body frame $\{B\}$ is given by the third column of $^B \mathbf{R}_U$. The upper arm axis will be used for determining the axis of rotation for internal/external rotation exercises.

V. MODULAR CONTROLLER

The modular control architecture implemented on the Robot Control Computer is shown in Figure 6. The exercise protocol is first parsed into a control mode based on the desired activation of configuration space variables. This code then determines which controller should be activated for the possible combinations of arm groups: scapula, shoulder, elbow pitch, and elbow orbit. These groups can implement impedance and admittance modes depending upon the availability of force sensing (required for admittance). In the case where both modes are feasible, e.g. rowing, the level of impedance often determines which mode will be implemented. Several of the control modules are discussed in more detail below.
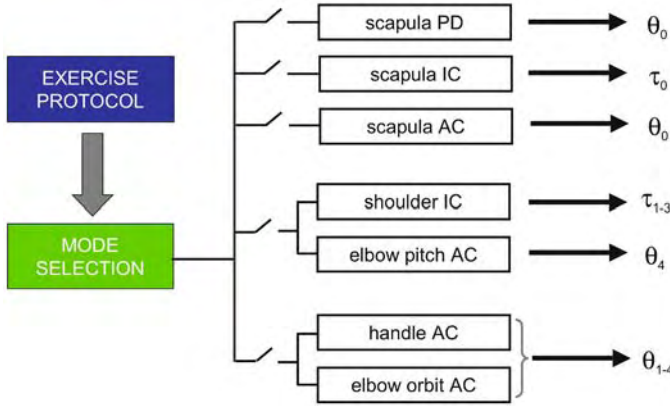

Fig. 6.   Control modules are activated by the mode selector depending upon the exercise protocol.

A. Shoulder-Elbow Admittance Module

The admittance controller shown in Figure 7 is used to convert the sensed contact forces at the hand and torque about the elbow orbital angle into desired movements of the exoskeleton [12]. Signals from the force-torque sensor at the hand are relayed to an admittance model of the virtual environment, which then outputs a desired velocity for the wrist, $\dot{p}_w$. In addition, a pair of uni-axial load cells mounted along the elbow pitch axis are used to determine the torque, τ_ϕ, exerted about the shoulder-wrist axis, p_w. The elbow torque is calculated by taking the product of the force output from the active load cell and multiplying it by the moment arm

$$\tau_\phi \quad \equiv \quad \mid \mathbf{p}_p \mid \mathbf{F}_e \qquad (11)$$

where $\mathbf{p}_p$ is the minimum distance from the elbow to the line SW, and F_e is the load cell output. The elbow roll torque is then integrated to produce a desired elbow orbit velocity $\dot{\phi}_d$ that is proportional to the torque about the p_w axis. The desired wrist and elbow orbit velocities are then converted to desired angular velocities at the joints using the inverse Jacobian, $J_{w\phi}$, and the resultant desired joint angles, $\dot{\theta}_d$, are then tracked using a proportional-derivative (PD) control law.
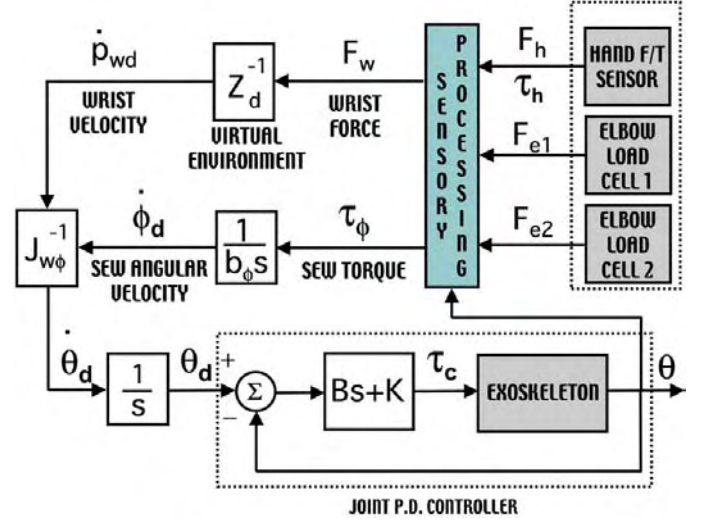

Fig. 7.   The admittance controller utilizes force inputs from the handle and elbow load cells to compute the desired wrist velocity and elbow orbit angular rate, which are then input to the inverse kinematics to determine the exoskeleton joint rates.

The inverse kinematics for the exoskeleton joint angles $[\theta_1 \ \theta_2 \ \theta_3 \ \theta_4]^T$ are found from $^0p_5$ and ϕ using the extended Jacobian approach [22]. Because of the complex dependence of the wrist position and, in particular, the elbow orbit angle on the joint angles, an analytical solution for the exoskeleton inverse kinematics is not realizable. Thus, a Newton-Raphson iterative procedure is used to determine the change in joint angles as a function of the desired change in wrist position and elbow orbit angle.

B. Shoulder Impedance Module

An impedance controller is used to attain the desired impedance about an arbitrary axis of the shoulder. The stiffness and damping about the desired Cartesian axes of rotation are set using the desired impedance Z_{GH_d}. The desired orientation of the glenhumeral (GH) joint is input to the controller as shown in Figure 8 and "differenced" with the sensed GH orientation computed from the forward kinematics of the shoulder joints to produce the angle-axis error $^B\Delta r_G$. The GH angular velocity error $^B\Delta\omega_G$ is then multiplied by the desired impedance Z_{GH_d} to produce the desired GH torque. F_d is then mapped to joint torques τ_d via the transpose-Jacobian, which is added to a feedforward compensation torque τ_{fwd} to produce the torque servo command τ_c.

VI. CONTROL EXPERIMENTS

The architecture of the control system is shown in Figure 9. The control station is used by the clinician to select protocols, initiate and terminate operation, and monitor subject data. The control station communicates over the Internet with the robot control computer, which is responsible for control of the robotic arm and overall patient safety. The arm controllers executing on the robot control computer produce
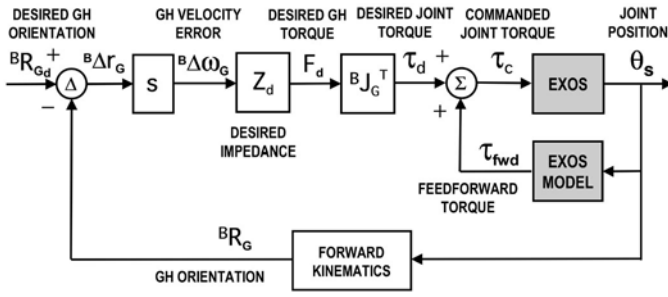
Fig. 8. Impedance controller used for shoulder axis rotation.

either a position or torque command, which is sent to the motor controller.

All safety decisions are carried out within the robot control computer and occur autonomously. The robot control computer runs a Linux real-time operating system, in order to guarantee meeting its safety deadlines. This setup allows for very high speed reaction by the computer safety system in the event of a component failure, communication error, or the patient attempting something they should not [23].

Controller algorithms that operate in Cartesian space run at 125 Hz, while joint-level controllers operate at 1000 Hz on a Galil 6-axis servoboard. A JR3 force-torque sensor attached to the gripper senses the forces and torques exerted by the patient, which are relayed to the robot control computer via a PCI board. Analog inputs from the elbow load cells are read by a separate National Instruments PCI A/D board.

Preliminary experiments were conducted to demonstrate the ability of the admittance controller to achieve desired impedances of the configuration space variables in translation of the handle and elbow orbit roll.
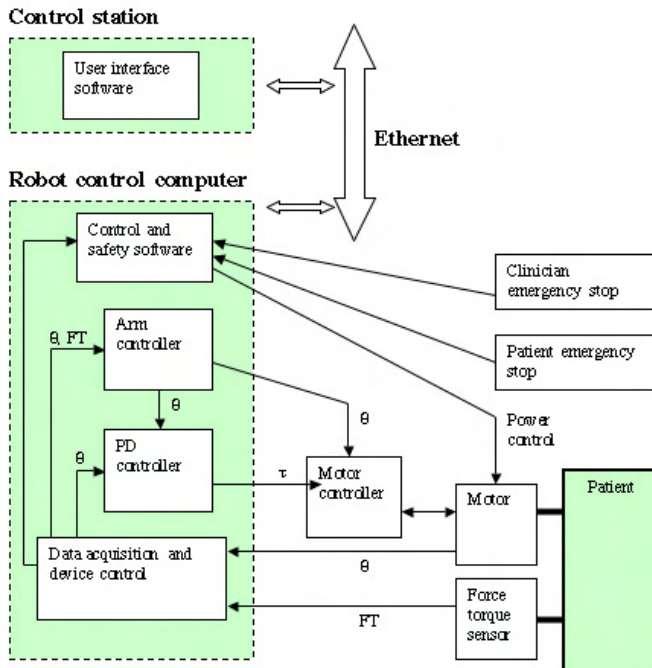
Fig. 9. Software control architecture.

A. Handle Translation with Spring-Damping

The damping and stiffness of the handle were set to 500 N/m/s and 500 N/m, respectively, and the elbow orbit angle was fixed. The initial joint position was approximately $\theta = [0° \ 90° \ -105° \ -90° \ 90° \ 0°]^T$ which corresponds to the upper arm straight down and the elbow flexed at $90°$ as seen in Figure 10.

Fig. 10. Configuration of the exoskeleton during experiments.

The position and sensed force on the handle in the x_0-direction during a forward push and "release" are shown in Figures 11 and 12, respectively. As the subject pushes the handle forward, the force builds up due to the restoration force in the virtual spring. After the subject reaches the maximum deflection, he lets the handle move back to the neutral position trying not to exert any force on the handle. This produces a classic, delayed exponential curve with a time constant of about 1 sec.

Fig. 11. Cartesian position of handle in x_0-direction for B=500 N/m/s and K= 500 N/m during push/release maneuver.

Fig. 12. Cartesian force on handle in x_0-direction during push/release test.
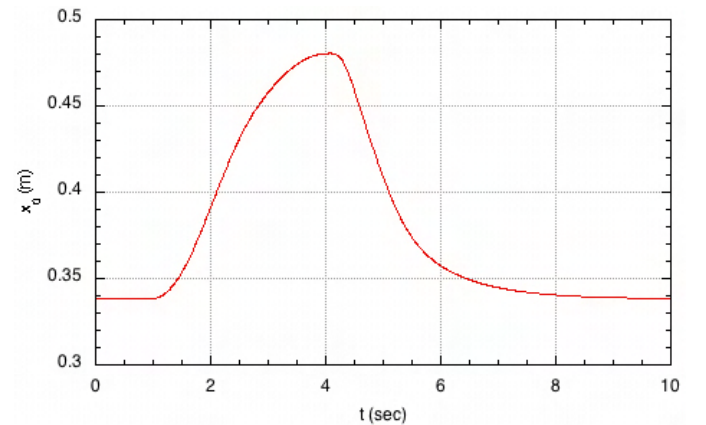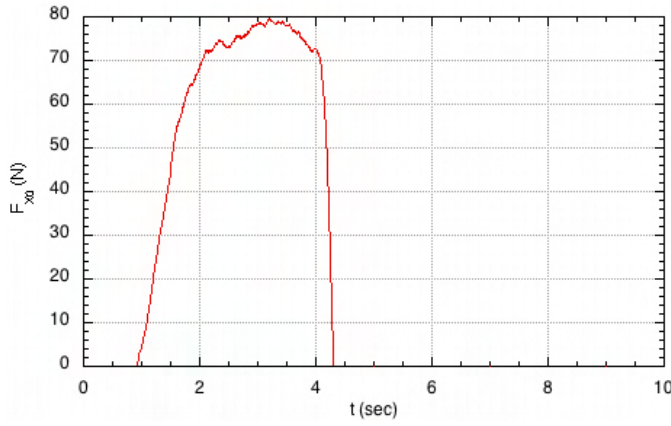

Fig. 14. Cartesian force on handle during square contour following task.

B. Handle Translation with Pure Damping

In this test, the subject attempted to trace out a square with the handle in the frontal plane (y-z base frame) using only dead reckoning. The damping of the handle was reduced to 250 N/m/s to allow freer movement, and the stiffness was set to zero so there was no restoration force. The elbow orbit angle was fixed, and the initial position was the same as the previous test.

The projection of the movement in the y-z base plane is shown in Figure 13. The subject starts at the bottom center of the plot and proceeds around the square clockwise when viewed from the front. The time history of the sensed handle forces in base frame coordinates are shown in Figure 14. F_y starts out negative which produces the leftward movement at the bottom of the plot, and then F_z becomes positive indicating the upward leg on the left side of the plot. Forces tend to remain fairly constant along the linear segments indicating constant velocity.

Fig. 13. Cartesian path of the handle in y-z plane for B=250 N/m/s during square contour following task.

C. Elbow Orbit Accommodation

In the next experiment, the subject attempts to execute a pure elbow orbit maneuver by "rolling" the elbow about the shoulder-wrist line in each direction at a constant velocity. The rotational damping b_ϕ was set to 50 N-m/rad/sec so that the rotational velocity is proportional to the exerted torque. The resulting elbow orbit angle and torque are shown in Figures 15 and 16, respectively. The torque alternates sign as the subject rotates the elbow first in one direction and then the other. The slope of the elbow orbit angle remains constant while the torque maintains the same value during the sequential rotations indicating a pure damping impedance.
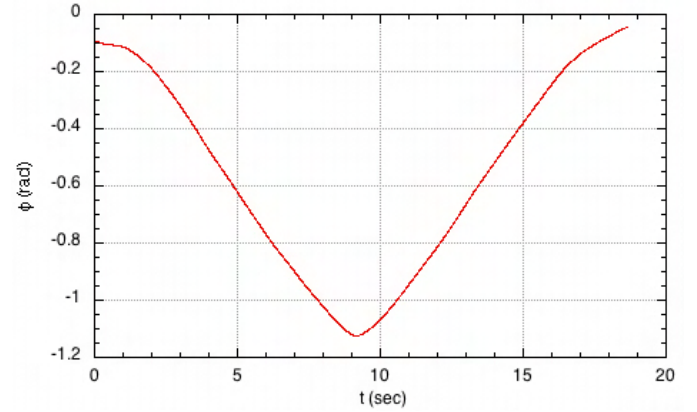
Fig. 15. Elbow orbit angle for B=250 N-m/rad/s during accommodation maneuver.

VII. Conclusion

A configuration-space approach was developed to decouple the control of the exoskeleton into modules that depend upon the specific rehabilitation protocol. The configuration space for a task consists of various combinations of anthropomorphic sets of joints which are mutually exclusive. Impedance and admittance control modes are generally available for each set, the choice depending upon the availability of force sensors and impedance gains.
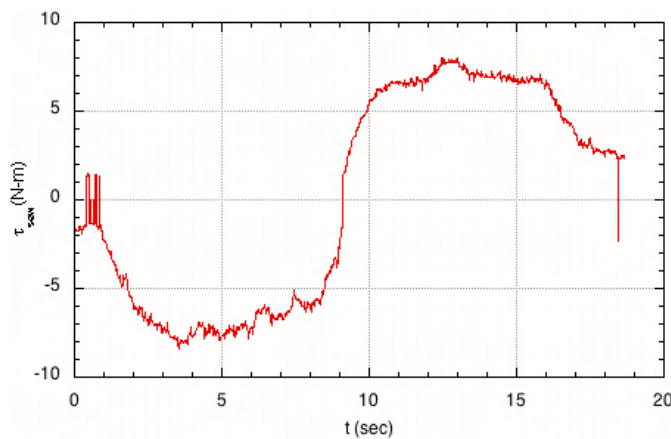
Fig. 16. Elbow axial torque during accommodation maneuver calculated from the elbow load cells.

An admittance controller was used to demonstrate control of handle forces and elbow orbit during a simulated task. A force-torque sensor at the handle was used to realize a haptic interface for simulating contour following with friction. Load cells mounted at the elbow were used to sense the torque about the elbow orbit axis and accommodate subject movement.

We are currently developing the impedance control module that will be used to generated iso-lateral rotational exercises about an arbitrary shoulder axis, where direct torque sensing is not possible. A feed forward model is needed to reduce the effect of the natural dynamics of exoskeleton. Initially, a lumped mass model will be developed for the gravity model, and joint level system identification will be used to estimate friction parameters.

In addition, admittance and PD tracking controllers are being developed for the scapula joint. A load cell is being integrated to the scapula lever arm that will sense upward forces applied by the subject's shoulder that can be used as force input to drive an admittance controller for shoulder shrug exercises. Alternatively, upper arm motion can be used to estimate shoulder elevation during arm maneuvers involving normal scapulo-humeral rhythm [24]. This modeled motion can be used as input to a PD controller to drive the shoulder joint and keep it aligned with the human GH joint.

Iso-lateral exercises are currently being developed that are close parallels to manual therapy and exercise machines such as shoulder flexion and rowing. From there, we will develop functional training protocols beginning with PNF patterns such as D1 and D2 and progressing toward simulated everyday tasks. The develop of clinical interfaces for choosing task parameters and patient monitoring is being done in parallel with the protocol development.

ACKNOWLEDGMENTS

REFERENCES

[1] D. G. Caldwell, C. Favede, and N. Tsagarakis, "Dextrous exploration of a virtual world for improved prototyping," in *Proc. of the IEEE Intl. Conf. on Robotics and Automation*, Leuven, Belgium, May 1998, pp. 298–303.
[2] L. Peterson and P. Renstrm, *Sports Injuries Their Prevention and Treatment, 3rd ed.* Champaign, Il.: Human Kinetics, 2001.
[3] P. Houglum, *Therapeutic Exercise for Athletic Injuries*. Champaign, Il.: Human Kinetics, 2001.
[4] A. Schiele and G. Visentin, "The esa human arm exoskeleton for space robotics telepresence," in *Proc. 7th Int. Symposium on Artificial Intelligence, Robotics and Automation in Space (iSAIRAS)*, Nara, Japan, 2006.
[5] H. Seraji, "Configuration control of redundant manipulators: Theory and implementation," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 4, Aug. 1989.
[6] G. C. Burdea, *Force And Touch Feedback For Virtual Reality*. New York: John Wiley and Sons, 1996.
[7] T. Massie and J. K. Salisbury, "The PHANTOM haptic interface: A device for probing virtual objects," in *Proc. ASME Winter Annual Meeting: Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, Nov. 1994.
[8] C. Carignan and K. Cleary, "Closed-loop force control for haptic simulation of virtual environments," *The Electronic Journal of Haptics Research (http://www.haptics-e.org)*, vol. 1, no. 2, Feb. 2000.
[9] M. Bergamasco, B. Allotta, L. Bosio, L. Ferretti, G. Parrini, G. Prisco, F. Salsedo, and G. Sartini, "An arm exoskeleton system for teleoperation and virtual environments applications," in *Proc. of the IEEE Intl. Conf. on Robotics and Automation*, San Diego, 1994, pp. 1449–1454.
[10] A. Frisoli, F. Rocchi, S. Marcheschi, A. Dettori, F. Salsedo, and M. Bergamasco, "A new force-feedback arm exoskeleton for haptic interaction in virtual environments," in *Proc. of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 2005.
[11] P. Swaim, C. Thompson, and P. Campbell, "The Charlotte intravehicular robot," NASA, Tech. Rep. N95-23703, 1995.
[12] C. Carignan and D. Akin, "Achieving impedance objectives in robot teleoperation," in *Proc. of the IEEE Conference on Robotics and Automation*, 1997, pp. 3487–3492.
[13] C. Clover, G. Luecke, J. Troy, and W. McNeely, "Dynamic simulation of virtual mechanisms with haptic feedback using industrial robotics equipment," in *Proc. of the IEEE Conference on Robotics and Automation*, Apr. 1997, pp. 724–730.
[14] D. Lawrence, "Impedance control stability properties in common implementations," in *Proc. IEEE Int. Conf. on Robotics and Automation*, 1988, pp. 1185–1190.
[15] A. Nakai, Y. Kunii, H. Hashimoto, and F. Harashima, "Arm type haptic human interface: Sensor arm," in *Intl. Conf. on Artificial Reality and Tele-Existence (ICAT)*, Tokyo, Japan, Dec. 1997, pp. 77–84.
[16] T. Nef, M. Mihelj, G. Colombo, and R. Riener, "Armin – robot for rehabilitation of the upper extremities," in *Proc. IEEE Int. Conf. on Robotics and Automation*, orlando, 2006, pp. 3152–3157.
[17] E. J. McCormick, Ed., *Human Factors Engineering, 3rd ed.* New York: McGraw-Hill, 1970.
[18] W. Prentice, *Rehabilitation Techniques in Sports Medicine, 3rd. ed.* New York: McGraw-Hill, 1999.
[19] M. Siff, *Supertraining, 6th ed.* Denver: Supertraining Institute, 2003.
[20] J. Craig, *Introduction to Robotics: Mechanics and Control, 2nd ed.* Reading, Mass.: Addison-Wesley, 1989.
[21] K. Kreutz-Delgado, M. Long, and H. Seraji, "Kinematic analysis of 7 dof manipulators," *Int. Journal of Robotics Research*, vol. 11, no. 5, pp. 469–481, 1992.
[22] C. Carignan and R. Howard, "A partitioned redudancy management scheme for an eight-joint revolute manipulator," *Journal of Robotic Systems*, vol. 17, no. 9, pp. 453–468, Sept. 2000.
[23] S. Roderick and C. Carignan, "An approach to designing software safety systems for rehabilitation robots," in *Proc. Int. Conf. on Rehabilitation Robotics (ICORR)*, Chicago, June 2005, pp. 252–257.
[24] T. B. Moeslund, C. B. Madsen, and E. Granum, "Modelling the 3d pose of a human arm and the shoulder complex utilising only two parameters," in *Proc. of the Intl. Conference on Model-based Imaging, Rendering, Image Analysis and Graphical Special Effects*, INRIA Rocquencourt, France, 2003.

8.2 Popa 2007: High Dynamic Range (HDR) …

Reprint begins on the next page and is eight pages.

High Dynamic Range (HDR) virtual bronchoscopy rendering for video tracking

Teo Popa [*], Jae Choi

Imaging Science and Information Systems (ISIS) Center, Department of Radiology,
Georgetown University, 2115 Wisconsin Avenue, Suite 603, Washington, DC, USA

ABSTRACT

In this paper, we present the design and implementation of a new rendering method based on high dynamic range (HDR) lighting and exposure control. This rendering method is applied to create video images for a 3D virtual bronchoscopy system. One of the main optical parameters of a bronchoscope's camera is the sensor exposure. The exposure adjustment is needed since the dynamic range of most digital video cameras is narrower than the high dynamic range of real scenes. The dynamic range of a camera is defined as the ratio of the brightest point of an image to the darkest point of the same image where details are present. In a video camera exposure is controlled by shutter speed and the lens aperture. To create the virtual bronchoscopic images, we first rendered a raw image in absolute units (luminance); then, we simulated exposure by mapping the computed values to the values appropriate for video-acquired images using a tone mapping operator. We generated several images with HDR and others with low dynamic range (LDR), and then compared their quality by applying them to a 2D/3D video-based tracking system. We conclude that images with HDR are closer to real bronchoscopy images than those with LDR, and thus, that HDR lighting can improve the accuracy of image-based tracking.

Keywords: virtual bronchoscopy, high dynamic range rendering, video-based tracking, tone mapping.

1. INTRODUCTION

In this paper, we present the design and implementation of a new rendering method that applies high dynamic range (HDR) lighting to a 3D virtual bronchoscopy system [1]. The two main optical parameters of a bronchoscope camera are the focal length and the exposure. The focal length determines the viewing angle and the size of the image, and the exposure limits the brightness of the image..Exposure is controlled in a camera by shutter speed and the lens aperture. Slower (longer) shutter speeds and greater (bigger) lens apertures produce greater exposures. Adjusting the aperture and the shutter speed of the camera makes it possible to deal with situations in which lighting conditions are far from optimal, such as situations where there are steep gradations of brightness, abrupt changes of illumination, and objects positioned against strong back light.

The bronchoscope camera relies on an internal light meter to determine proper exposure. Given an ISO sensitivity (ISO 100 for bronchoscopes), the meter will determine the camera settings (aperture f-number and shutter speed) that gives the best picture. To determine the best picture, the meters are calibrated to determine an exposure which will produce a picture that is, on average, grey. Therefore, a more stable lighting and rendering model is required for virtual bronchoscopy to simulate real camera settings.

---

[*] E-mail:popa@isis.georgetown.edu

This rendering model is computed in two steps. The first step is to compute the luminance values in high dynamic range (HDR) space [2], and the second step is to map the computed values to the values appropriate for display using a tone mapping operator that is common in computer graphics. The tone mapping operator is the computer graphics equivalent of bronchoscope exposure and ensure that the virtual bronchoscope image luminance will be grey on average.

When the rendering is done in HDR space, the camera settings can be adjusted and an accurate and realistic rendering can be achieved. By having virtual images closer to video images, it is possible to improve the accuracy and speed of a video-based tracking method (by improving the speed of convergence of the optimizer).

Video-based tracking of the bronchoscope can be achieved by 2D/3D registration using an optimizer that will search for the best matching of the video image to the virtual image from a 3D model of the bronchial tree. In the ideal case, the virtual images should match perfectly the video images and they should be generated in real time for practical usage of video-based tracking.

2. MATERIALS AND METHODS

Bronchoscopy is a procedure that allows physicians to look at patient's airways through a thin flexible viewing instrument called a bronchoscope. Current bronchoscopy systems are using video bronchoscopes, which incorporate a CCD sensor at the distal tip of the bronchoscope, replacing the fragile fiberoptic system used in earlier devices. A video processing unit provides high resolution color images. A typical video bronchoscope is shown in Figure 1.

Figure 1. Pentax 1530-T2 flexible video bronchoscope used in this study.

The bronchoscopy system that we used is a Pentax 1530-T2 flexible video bronchoscope with a focal range of 3~50mm, a shutter speed between 1/60-1/200 sec, and a viewing angle of 120 degrees. It can also visualize high-quality sharp images captured by a 410K pixel CCD sight spot (with a ISO 100 sensitivity). It also automatically determines the correct exposure of the scene. The lighting is generated from a mains-powered fiber-optic bronchoscope with a 300-watt white light lamp (Xenon). This light source emits light from the entire visible spectrum.

To acquire the images from the bronchoscope we used a Matrox Meteor II acquisition card. Images were saved in raw format, and later converted from RGB (red, green, blue) to gray-scale .There is no exact conversion because of the varying sensitivity response curve of the bronchoscope camera sensor to light as a function of wavelength. Therefore, we used the following common conversion:

$$L_w(x, y) = 0.3 * R(x, y) + 0.59 * G(x, y) + 0.11 * B(x, y) \quad (1)$$

where $L_w(x, y)$ is the "world" luminance for pixel (x, y), and R(x, y), G(x, y), B(x, y) are the color values.

To create the virtual bronchoscopic images, we first rendered a raw image in absolute units (luminance); then, we mapped the computed values to the values appropriate for video-acquired images. When absolute units are required, so are the exact proprieties for light sources and materials. These properties are hard to get or measure, and sometimes the tolerance range of a given property is too high [3]. Data for artificial light sources can be obtained from the bronchoscopy manufacturers.

The proprieties needed to define BRDFs (Bidirectional Reflectance Distribution Function) for anatomical materials used in the scene is far more difficult to find than the light source data [5]. For the first step, we used the surface rendering because high rendering speed is needed to achieve interactive video tracking. To create the surface patch, the Marching Cubes algorithm was implemented. To reduce the number of triangles within the surface patch, a decimation algorithm was used.

To emulate the exposure and shutter speed, we used the tone mapping operator by Reinhard [6], which maps a high dynamic range image space into a low dynamic range space suitable for screen display. We applied this scaling operator that is analogous to setting exposure in a camera. We set the tonal range of the output image based on the scene's key value. We considered the log-average luminance as a useful approximation to the key of the scene. This quantity is computed as follows:

$$\overline{L}_w = \frac{1}{N} \exp\left(\sum_{x,y} \log(\delta + L_w(x, y))\right) \quad (2)$$

where $L_w(x, y)$ is the "world" log average luminance for pixel (x, y), N is the total number of pixels in the image and δ is a small value to avoid the singularity that occurs when black pixels are present in the image.

After the average scene luminance is calculated, the HDR scene texture is scaled based on a target average luminance calculated as follows:

$$L(x, y) = \frac{a}{\overline{L}_w} L_w(x, y) \quad (3)$$

where $L(x, y)$ is a scaled luminance and $a = 0.18$ (because the scene has the normal key value, we map this to middle-grey of the displayed image).

For low-key or high-key images, it is possible to map the log average to different values of a (typically a can vary from 0.18 up to 0.36 and 0.72 and vary down to 0.09 and 0.045). The value of parameter a is called the "key value", because it relates to the key of the image after applying the above scaling. The key of a scene indicates whether it is subjectively light, normal, or dark. A white-painted room would be high-key, and a dim stable would be low-key. The main problem in Equation 3 is that many scenes have predominantly a normal dynamic range, but have a few high luminance regions

near highlights. This issue can be addressed by a different simple tone operator that compresses mainly the high luminances [8] [9]:

$$L_d(x, y) = \frac{L(x, y)}{1 + L(x, y)} \qquad (4)$$

which brings all luminance within a displayable range. The high luminances are scaled by approximately $\frac{1}{L}$, while the low luminances are scaled by 1. The denominator causes a graceful blend between these two scalings.

In addition, the areas of the video image where information is lost due to extreme brightness are described as having glare highlights. We simulate these effects by using a bloom effect. If absolute units for the light are used during the render process, these effects can be produced after the virtual image has been rendered. The bloom effect is achieved by blurring the image, and drawing it over top of the original image.

Graphics libraries such as OpenGL and DirectX contain a shader programming model in which per-pixel and per-vertex operations can be done using the hardware. For real-time rendering, we implemented the HDR rendering on a programmable GPU based on OpenGL. Current graphics cards have programmable vertex and fragment shaders that give the programmer the choice of lighting models and shading.

| (a) Real video image | (b) Generated virtual HDR image from 3D CT |
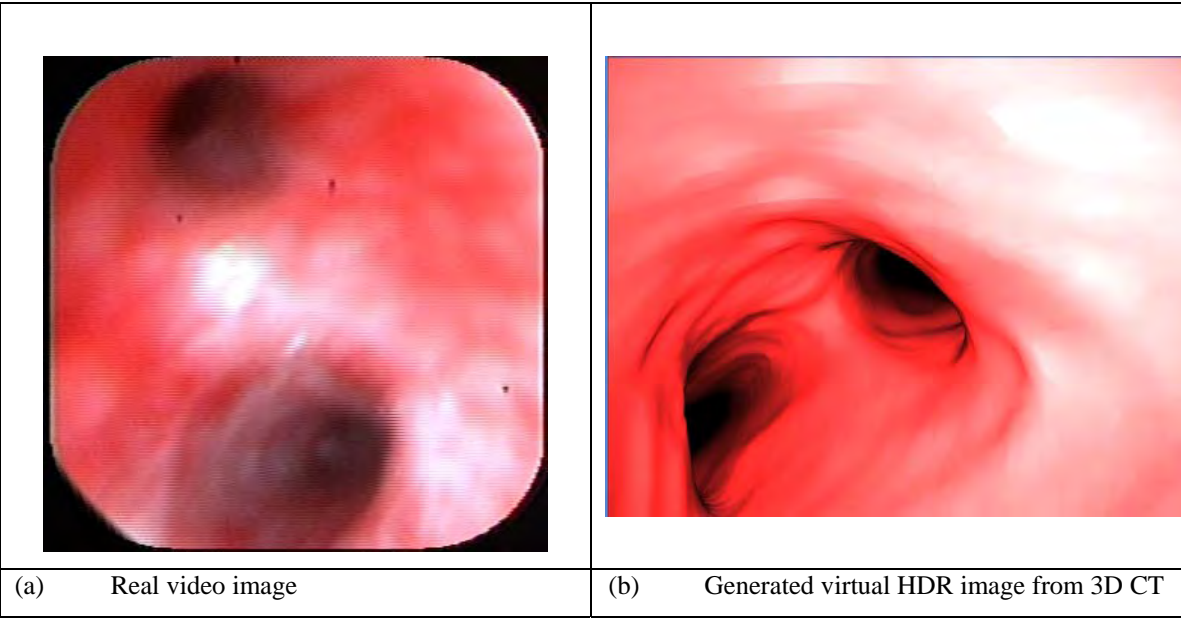
Figure 2: Example of image sources.

We used the *EXT_framebuffer_object* extension from OpenGL for its render-to-texture operations. It renders a scene into a floating point buffer using a phong lighting model (ambient, diffuse and specular lighting were calculated) and simple shadow mapping. For the lighting model, we used spotlights and the spherical mapping to replace the specular highlights calculations by looking up these values in the sphere map. The HDR portion of the image was then calculated through a separable Gaussian blur filter and composited back onto the low dynamic range (LDR) portion, resulting in a realistic glow effect.

| (a) Low Dynamic Range lighting* | (b) High Dynamic Range lighting* | (c) Real video image* |
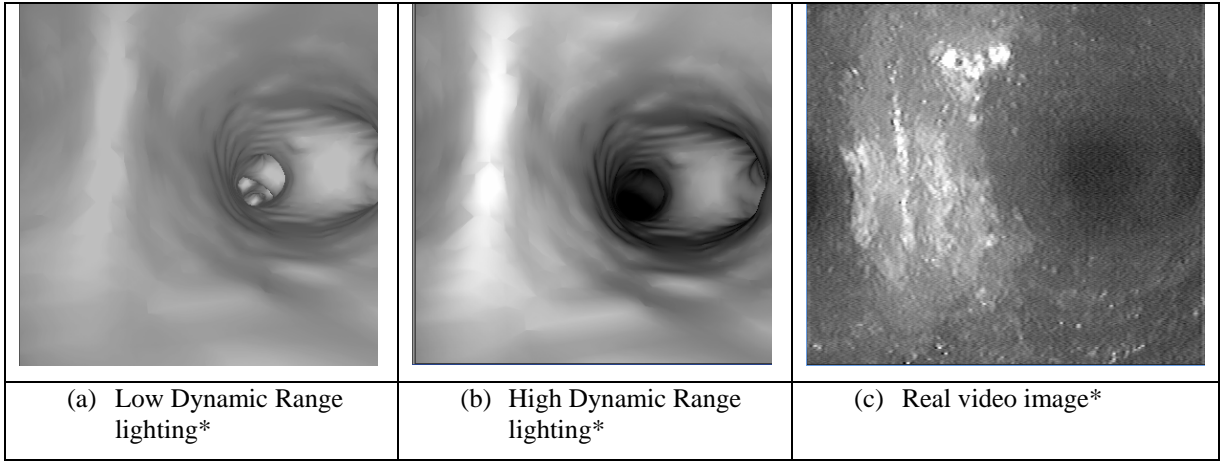
*The gray values represent the luminace intensities.

Figure 3. Typical HDR, LDR and video images captured from our software application

Figures 2 and 3 show the generated virtual images with different renderings and the acquired video image from the bronchoscope, in gray scale and with color. We used the generated HDR virtual bronchoscopy images to track the video bronchoscope camera position inside pre-acquired 3D CT images. Tracking was implemented using a 2D/3D image-based registration technique.

A typical 2D/3D image-based registration contains a optimization module, as well as a metric between virtual and video images. The most often used metric in image-guided interventions is the mean squared error. Unfortunately, it does not correspond to human perception, and sometimes images that look similar can be more different than obviously different images when the mean square metric is used. But, by improving the photorealism in the synthesis of computer-generated imagery, we are a step closer to achieving registration based on the mean squared error measure.

To overcome the problems related to mean squared error metric sensitivity we used a more robust metric called normalized mutual information, which requires additional computational time [10]. The normalized mutual information between two images A and B can be computed by:

$$NMI = \frac{H(A) + H(B)}{H(A, B)} \qquad (5)$$

where $H(A,)$, $H(B)$ and $H(A,B)$ represents image entropies and can be easily computed from the histograms of the images.

3. RESULTS

For our bronchoscope, we have assumed, based on the work of Skilton [4], that the brightness (luminance) of the light spot in our rendering method is 2000 cd/m^2. For hardware acceleration we used an Nvidia Geforce 6800 GT video card, which has implemented Shader Model 3.0. To validate that HDR rendering has advantages over normal lighting conditions, we performed one experiment. In this experiment, we acquired video images from a bronchoscope at five positions. We also recorded the position of the tip of the bronchoscope with an Aurora electromagnetic (EM) tracking system (Northern Digital Inc, http://www.ndigital.com). After these steps, we performed fiducial registration that gave the transformation between CT image space and EM space. We generated the virtual images at the tip position. This virtual image was used as a starting position for 2D/3D registration. In this way, we could compare video tracking with HDR virtual-based tracking and LDR virtual-based tracking. After finding a good match between the virtual images and the video images for a particular position, we tested the sensitivity of the mutual information metric between HDR and video images, and also between LDR and video images.

We changed the position of the virtual camera sequentially in the x, y, and z directions by a step of 8 mm. We computed the mutual information value between each generated image and the corresponding video image. If the generated images were closer to the video images, the mutual information value increased (Figure 4).

(a)   Translation in x direction

(b)   Translation in y direction
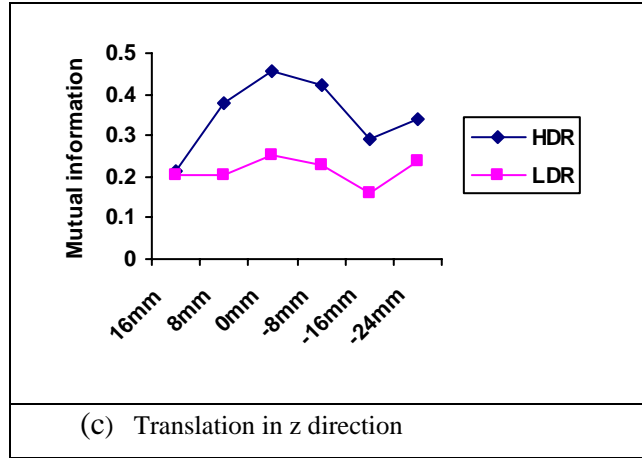
(c) Translation in z direction
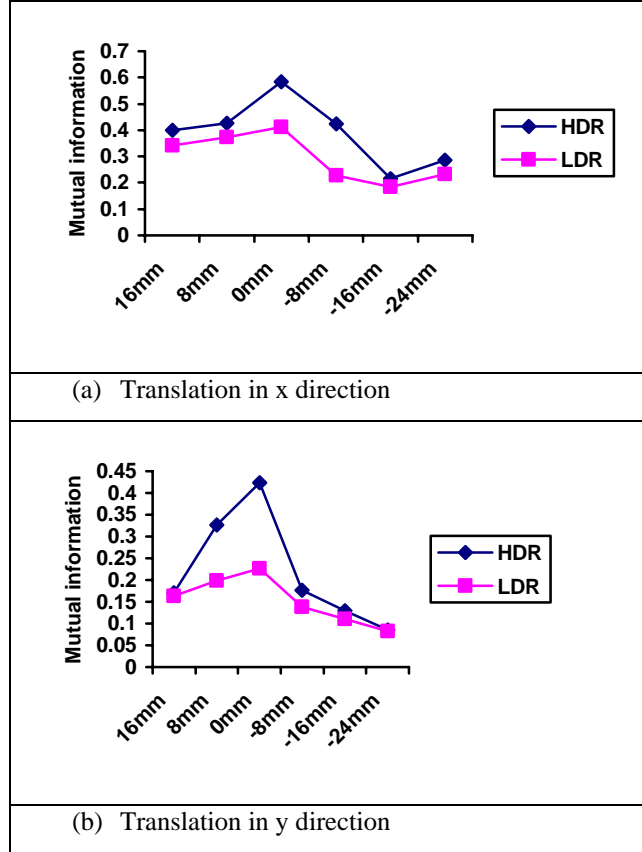
Figure 4. Translation in each direction along the best fit video image

Figure 4 shows the results. The values on the horizontal axis of the three graphs(a), (b) and (c) represent the distance in mm from the virtual camera position that matches the video camera position. The values on the vertical axis represent the normalized mutual information value. From Figure 4, it can be observed that when the distance of the virtual camera increases, the mutual information value becomes lower, which means the distant virtual image will match less with the video image. The mutual information computed for HDR generally had a larger value than mutual information for LDR, which shows a better fitting of the virtual image to the video image.

4. CONCLUSIONS

High dynamic range visualization allowed us to generate images that are better suited for 2D to 3D registration. By considering camera exposure simulated though tone mapping, it is possible to compensate for different bronchial exposures. In addition, it is possible to simulate auto exposure by extending the tone-mapping calculations. We currently tested the similarity between HDR images and real bronchoscopy with mutual information.

However, mutual information calculation is computationally expensive, and limits registration performance even on high-end machines. By refining the HDR rendering, we plan to replace this cost function with the sum of squared differences. Additionally, we intend to produce our own HDR renderings based on real measured material, surface and light sources. Current bronchoscope systems use relatively simple exposure control, but we foresee that, with the advancements of bronchoscopic image quality, a HDR based rendering model will be better suited.

ACKNOWLEDGEMENTS

REFERENCES

1. Debevec, P., and Malik, J.. Recovering High Dynamic Range Radiance Maps from Photographs, ACM Computer Graphics, SIGGRAPH 1997 Proceedings, 31(4):369-378.

2. DiCarlo, J.M., and B.A. Wandell, Rendering High Dynamic Range Images, Proceedings of the SPIE: Image Sensors, 3965:189-198.

3. Matkovic K. , Tone Mapping Techniques and Color Image Difference in Global Illumination, 1997, Institute of Computer Graphics and Algorithms, Vienna.

4. Skilton H; Parry; Arthurs G. J. Hiles P.; A study of the brightness of laryngoscope light. Anaesthesia , 1996, vol. 51, No. 7, pp. 667-672.

5. Chung, A.J., Deligianni, F., Shah, P., Wells, A., Yang, G.: Patient-specific bronchoscopy visualization through BRDF estimation and disocclusion correction. IEEE Trans. Med. Imaging 25(4): 503-513 (2006).

6. Reinhard, E., M. Stark, P. Shirley, and J. Ferwerda, Photographic Tone Reproduction for Digital Images, ACM Transactions on Graphics, 21(3):267-276.

7. Ashdown, I., Radiosity: A Programmer's Perspective, 1994, John Wiley & Sons, New York, NY.

8. Strobel, L., Compton, J.,Current C, I., and Zakia, R. 2000. Basic photographic materials and processes, second ed. Focal Press.

9. Mitchell, E. N. 1984. Photographic Science. John Wiley and Sons, New York.

10. Hajnal, J., Hawkes, D. J., and Hill, D.. Medical Image Registration. CRC Press, 2001.

8.3 Stenzel 2006: Augmented Lung Biopsy …

Reprint begins on the next page and is one page.

Augmented Lung Biopsy under CT Fluoroscopy:

Workflow and Registration

R. Stenzel[a], F. Banovac[a], K. Cleary[a]

[a] *Imaging Science & Information Systems (ISIS) Center, Department of Radiology, Georgetown University, Medical Center, Washington, DC, USA*

Keywords: lung; biopsy; segmentation; registration; deformable; elastic; rigid; CT; CT-Fluoroscopy

1. Introduction

One method for lung cancer diagnosis is a CT-fluoroscopy (CTF)-guided lung biopsy. During the intervention the physician is guided by repeated axial CTF images. The problem of mentally reconstructing the 3D scene from a single CTF image can be a difficult task for the physician. Therefore, we have been investigating methods for providing 3D image guidance to the physician during this procedure as described here.

2. Methods

Our workflow starts with the acquisition of a pre-operative CT volume at breath-hold. The entry and target points are then defined by the phyisican using the software package Image-Guided Lung Biopsy (IGLungBiopsy) developed by our group. The software also includes a semi-automatic lung segmentation capability. The physician then acquires several intra-operative CTF images at breath-hold showing the needle's cross sections and skin fiducials which were placed at the beginning of the procedure. The needle is segmented and the CTF images are registered to the CT volume via rigid and deformable registration using Thirion's demons [1] as described in [2]. The aim is to compute a deformation field between the CTF images and the CT volume to calculate the needle's position within the CT volume. Finally, a 3D rendering showing the CT volume, needle, and planned path is provided to the physician for guidance.

3. Results

Several CT volumes and CTF images were acquired during an approved swine study. The evaluation of the rigid registration gave a mean error of 0.2mm and a std. deviation of 0.24mm. The evaluation of the deformable registration was done using a known deformation field as ground truth. The deformation field (mean deformation < 2.7mm) was applied to the CT volume and a subvolume of the deformed volume was registered to the original CT volume. The mean error (x, y, z) was < 0.9mm, mean error (only z) was < 0.5mm, and the std. deviation (x, y, z) was < 0.6mm.

4. Conclusion

We are developing a new workflow for lung biopsies and techniques for registration and visualization to provide the physician a visual guidance during a lung biopsy. Our results so far in subvolume to volume registration using deformable registrations showed good agreement.

References

[1] J. Thirion. *Image matching as a diffusion process: an analogy with Maxwell's demons*. Medical Image Analysis 2(3): 243-60, Oxford University Press, 1998.

[2] Z. Yaniv, R. Stenzel, F. Banovac, K. Cleary. *A Realistic Simulation Framework for Assessing Deformable Slice-to-Volume (CT-Fluoroscopy/CT) Registration*. Medical Imaging 2006, Visualization, Image-Guided Procedures, and Display, Proc. SPIE Vol. 6141, p. 116-123.

8.4 Stenzel 2007: Precision Instrument Placement …

Reprint begins on the next page and is eight pages.

Precision Instrument Placement Using a 4-DOF Robot with Integrated Fiducials for Minimally Invasive Interventions

Roland Stenzel [*a], Ralph Lin [a], Peng Cheng [a], Gernot Kronreif [b], Martin Kornfeld [b], David Lindisch [a], Bradford J. Wood [c], Anand Viswanathan [c], Kevin Cleary [a]

[a] Computer Aided Interventions and Medical Robotics (CAIMR), Imaging Science & Information Systems (ISIS) Center, Department of Radiology, Georgetown University Medical Center, Washington, DC, USA
[b] ARC Seibersdorf Research GmbH, Austria
[c] Diagnostic Radiology Department, NIH Clinical Center, Bethesda, MD, USA

ABSTRACT

Minimally invasive procedures are increasingly attractive to patients and medical personnel because they can reduce operative trauma, recovery times, and overall costs. However, during these procedures, the physician has a very limited view of the interventional field and the exact position of surgical instruments. We present an image-guided platform for precision placement of surgical instruments based upon a small four degree-of-freedom robot (B-RobII; ARC Seibersdorf Research GmbH, Vienna, Austria). This platform includes a custom instrument guide with an integrated spiral fiducial pattern as the robot's end-effector, and it uses intra-operative computed tomography (CT) to register the robot to the patient directly before the intervention. The physician can then use a graphical user interface (GUI) to select a path for percutaneous access, and the robot will automatically align the instrument guide along this path. Potential anatomical targets include the liver, kidney, prostate, and spine. This paper describes the robotic platform, workflow, software, and algorithms used by the system. To demonstrate the algorithmic accuracy and suitability of the custom instrument guide, we also present results from experiments as well as estimates of the maximum error between target and instrument tip.

Keywords: instrument guide, robot, minimally invasive interventions, registration, segmentation, fiducial, CT, path planning, image-guided therapy, medical robotics

1. INTRODUCTION

Minimally invasive procedures are increasingly attractive to patients and medical personnel because they can reduce operative trauma, recovery times, and overall costs. However, during these procedures, the physician has a very limited view of the interventional field and the exact position of surgical instruments. Moreover, the intervention has to be performed through small openings without direct access to the anatomy. Consequently, these interventions can be very difficult to perform, and not all physicians are trained in these techniques. For instance, only two percent of all physicians trained to perform conventional, invasive prostate cancer surgery are also trained to perform its minimally invasive counterpart [1].

To address these drawbacks, we have created a new platform for precision placement of surgical instruments during minimally invasive interventions, based upon a small four degree-of-freedom robot (B-RobII; ARC Seibersdorf Research GmbH, Vienna, Austria) [3]. This platform consists of three components: a custom-designed instrument guide with a built-in fiducial pattern, a precision placement robot, and control software. The control software automatically extracts the locations of the fiducials from a pre-operative or intra-operative CT scan and registers the robot coordinate system to the CT coordinate system [2]. After this step, the robot orients the surgical instrument guide along the physician-selected path and places the instrument at the desired skin entry point. This paper presents the platform design, as well as our assessment of the system's accuracy.

*stenzel@isis.georgetown.edu; www.isis.georgetown.edu

2. METHODS

Our new compact instrument guide contains a spiral pattern of nineteen 1mm fiducials (Tantalum Beads; Tilly Medical, Lund, Sweden) (Figures 1, 2). This guide is designed to be used with either a coaxial radio-frequency ablation probe (LeVeen System; Boston Scientific, Natick, Massachusetts, USA) or a triple cluster radio-frequency ablation probe (Cool-Tip RF™; ValleyLab, Boulder, Colorado, USA).

Figure 1. Custom instrument guide with an integrated spiral fiducial pattern.
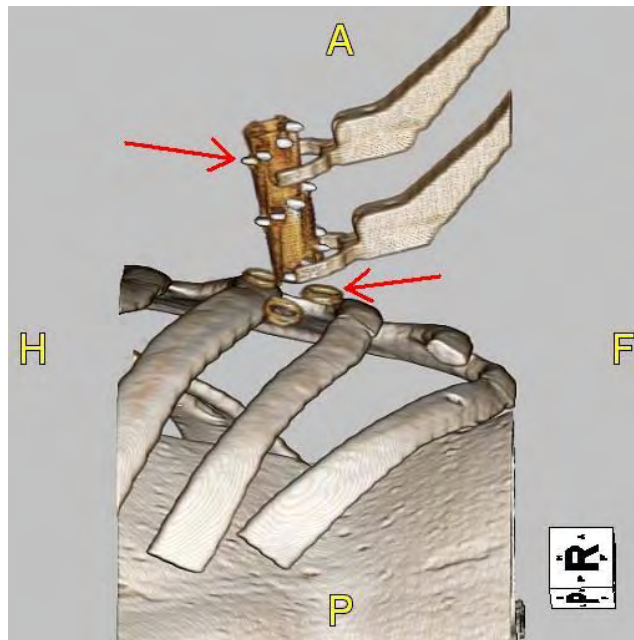
Figure 2. TeraRecon software reconstruction of CT scan of the robot's fingers, instrument guide, 1mm fiducials (which appear as white ellipsoids - see left arrow), and phantom. The donut fiducials (see right arrow) were used only for a fast visual inspection of the guide's position.

The guide is coupled to the B-RobII robot via two carbon fiber fingers, each of which is attached to one 2-degree-of-freedom module (Figure 3). The modules can be shifted within the working area (translation in x and y). Together with the instrument guide, they establish a parallelogram kinematics, which results in a rotation of the instrument guide around x and y (remote center of motion). The robot is highly portable, can be set up quickly and easily, and has an easily adaptable custom Graphical User Interface (Figure 4).

Figure 3. B-RobII precision placement modules (two degrees of freedom each).

Figure 4. Graphical User Interface showing axial, coronal, sagittal, and 3D rendered views of the data. All views include an overlay of the instrument guide (yellow cylinder), the instrument (green tube), and the working area (violet square).
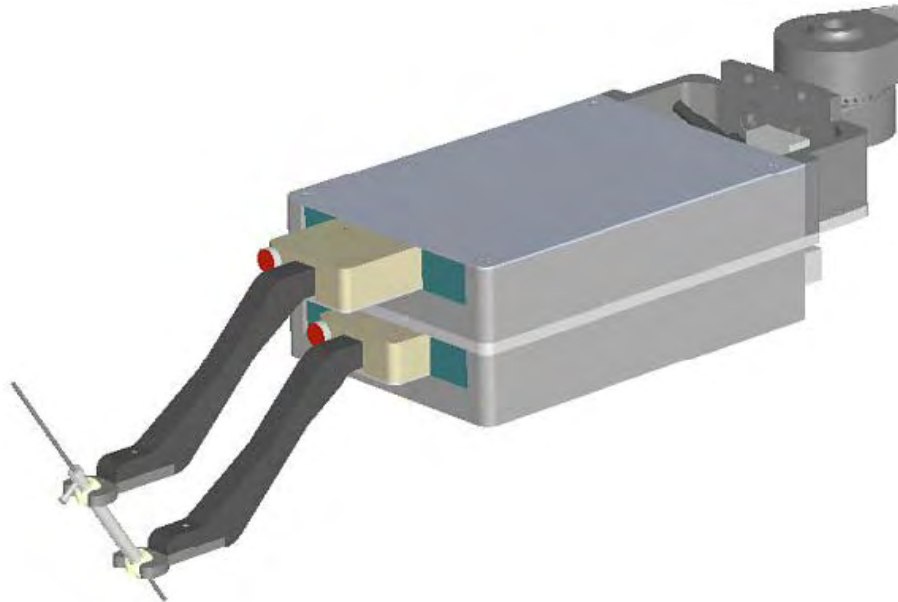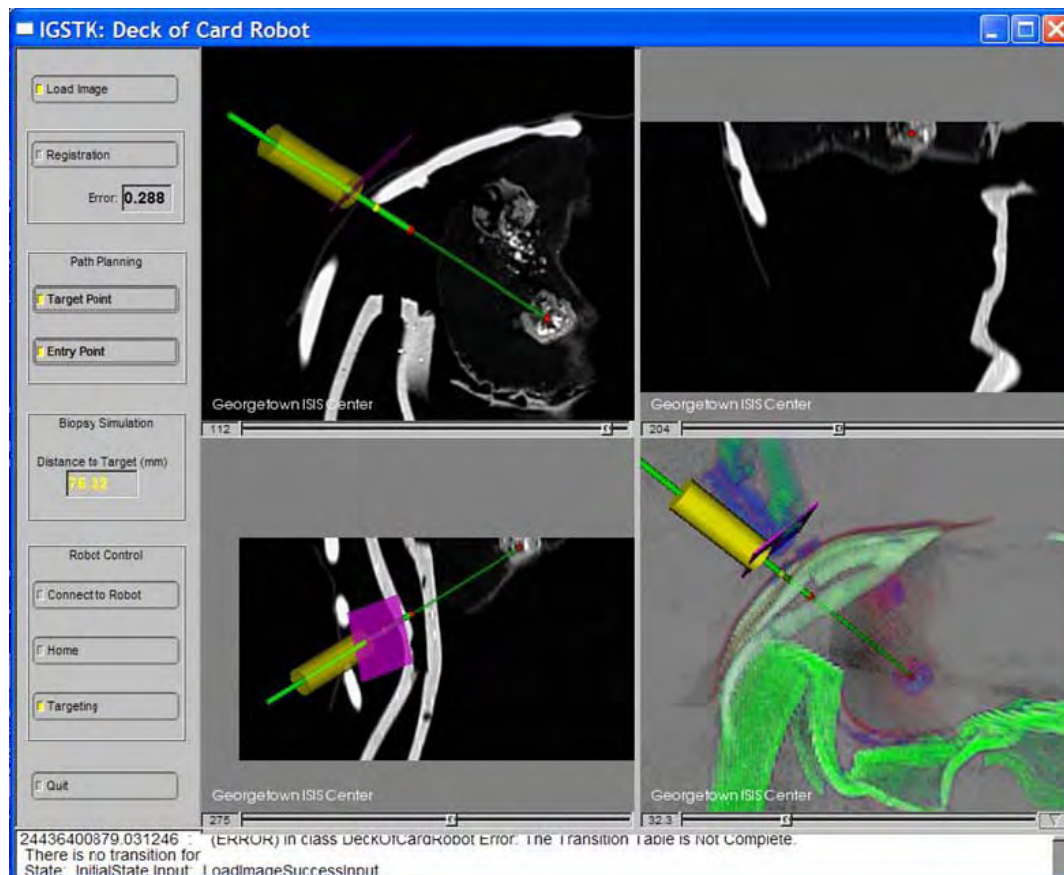
The scenario envisioned for robot operation in the clinical environment is as follows (Figure 5). First, the robot control box and intervention planning computer are set up prior to the patient's arrival. Next, the patient is placed on the CT table. Then, the robot, including the instrument guide, is mounted on the frame and positioned so that the robot's working area is close to the assumed skin entry point. After the robot is positioned as such, a set of axial CT images (5cm along the table) is obtained. The setup from a phantom validation study is shown in Figure 6.
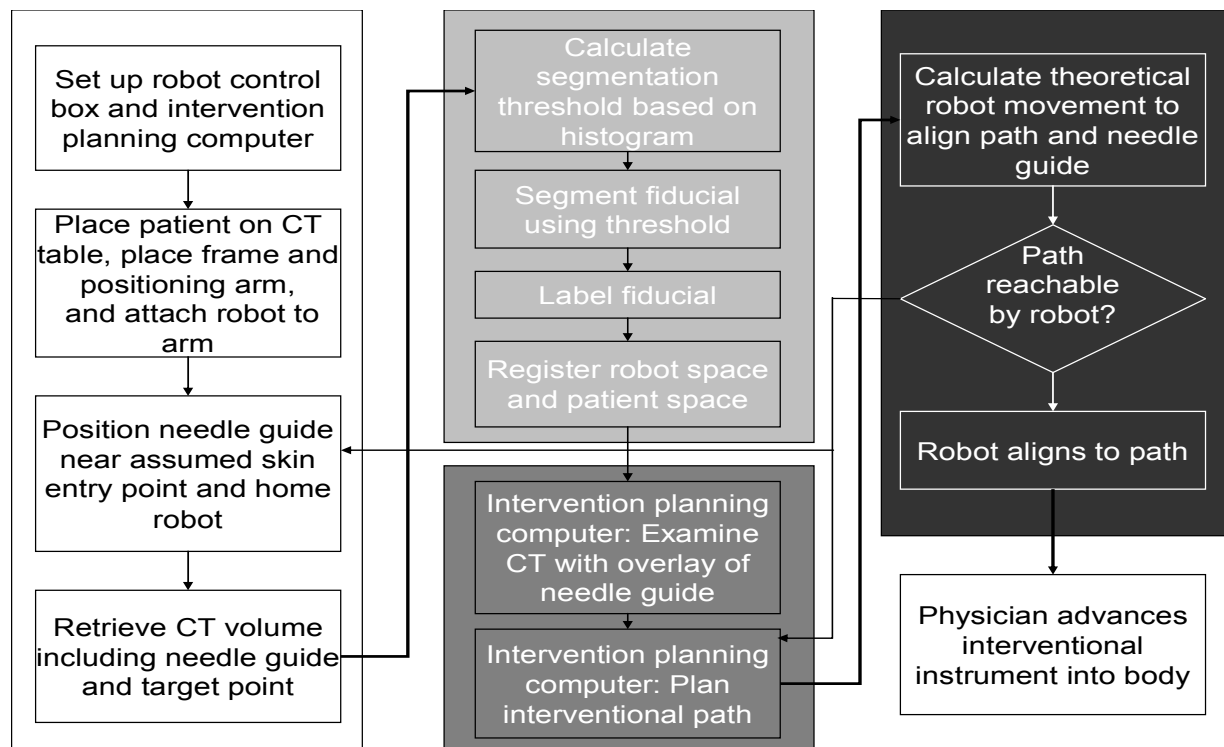

Figure 5. Proposed clinical workflow.

The set of CT images is loaded into the robot control software and is then used to automatically register the robot and the patient. First, an intensity threshold is calculated using the CT's histogram. This threshold is defined by a local minimum close to the upper end of the histogram. Then, all voxels above the threshold are segmented and labeled. The labeling algorithm assigns all voxels that are part of the segmentation result to a consecutively numbered fiducial area. Up to now, the relationship between the segmented fiducial areas and the fiducials in the instrument guide's model is unknown. However, the greatest variance by any projection of the fiducials is the same for the segmented fiducial areas and for the fiducials in the instrument guide's model. Therefore, Principal Component Analysis (PCA) is used to align the set of fiducial areas and the set of fiducials in the instrument guide's model. Next, all fiducial areas are sorted according to the real fiducials positions in the known fiducial pattern using the values of the first principal component. The result is a table that represents the one-to-one relationship between the fiducials in the instrument guide's model and the fiducial areas appearing in the CT. Then, the actual transformation between the CT (patient) space and the robot space is calculated using a closed-solution, paired-point registration [2].

Once the registration is performed, the physician can examine the CT images using the graphical user interface. The visualization now includes a virtual representation of the instrument, the instrument guide, and the robot's working space at its current position (Figure 4). The physician then selects the skin entry and target points, avoiding any obstacles such as ribs, major organs, or vessels. The robot aligns the instrument guide along the chosen path, and the physician performs the intervention by advancing the instrument into the body along the path. The physician intervenes only to

select the skin entry and target points within the CT scan; thus the interaction between physician and the application is minimized. The four quadrant display in the graphical user interface was created using the Image Guided Software Toolkit (IGSTK) [4, 5], an open-source software program, whose development was led by our group.
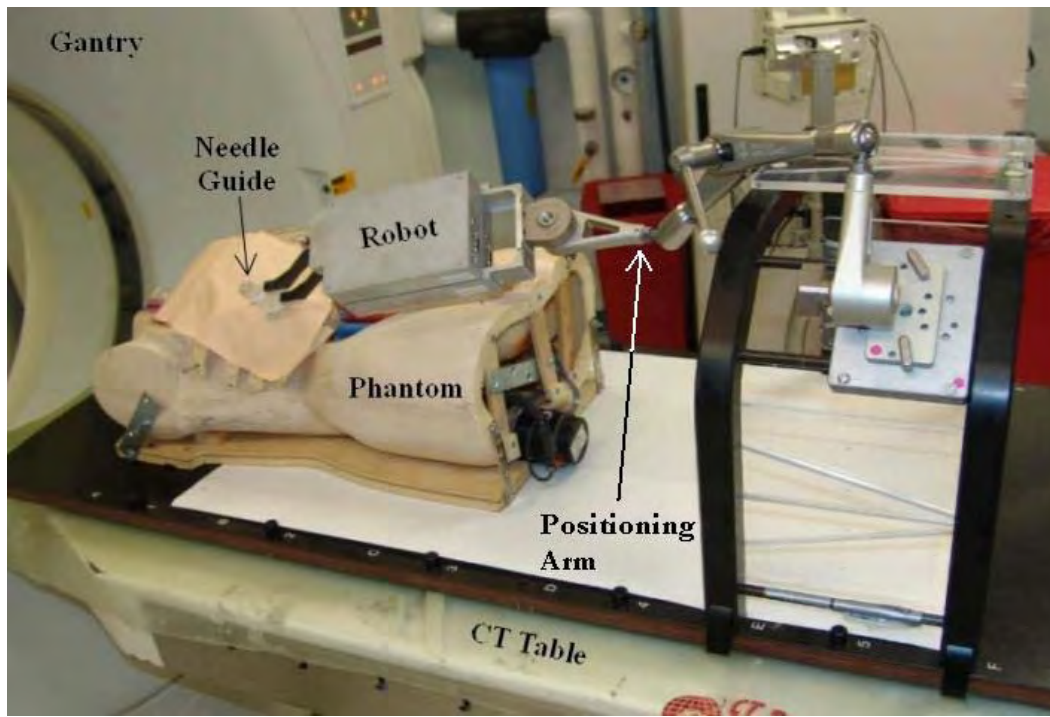

Figure 6. Robot setup in the CT suite, with phantom.

3. EXPERIMENTAL SETUP

To evaluate the integrated system's accuracy, we used a modified workflow**.** We developed a novel accuracy test that uses the robot itself to evaluate the registration algorithm. The error calculation therefore is free from external influences and external errors, like measurements by rulers or optical tracking systems. The experimental setup uses only a 15G biopsy needle as an end-effector and several CT scans from different robot positions within the working area. Thus, we performed the experiment using the setup shown in Figure 6, but without a patient or a phantom. As shown in Figure 7, we divided the evaluation into data gathering and data analysis steps.

Data gathering. We placed the robot in a random location using the positioning arm, homed the robot, and took several CT data sets as follows. The first CT (CT_H) was taken immediately after homing the robot. Then we created a list of three random instrument guide positions NGP_G (translation and rotation in x and y directions). The intervention planning computer then commanded the instrument guide (holding a 15G needle) to position n from NGP_G, and a CT (CT_n) was taken. This step was repeated for all positions on the list. After CT scans had been taken for all positions from NGP_G the entire process was repeated twice without any changes to obtain more statistically significant results.

Data analysis. The error was calculated as follows. For every CT_n , the needle was segmented, the needle center was calculated, and four center points – two at each end of the needle – were combined into four point sets $P_{1..4}$ (each with two points from each end). CT_H was then registered to the robot as described in Section 2 to obtain the transformation between robot space and CT (patient) space. The point set P_i was used for defining the skin entry and target points. The intervention planning computer then calculated the desired instrument guide position to align with the point set P_i , and the average instrument guide position for all $P_{1..4}$ was calculated. This average instrument guide position became position n in the list of calculated instrument guide positions NGP_C. The difference between NGP_C and NGP_G is the error, because NGP_G (ground truth) was used to get CT_n , and NGP_C is in turn based on points $P_{1..4}$ from CT_n.
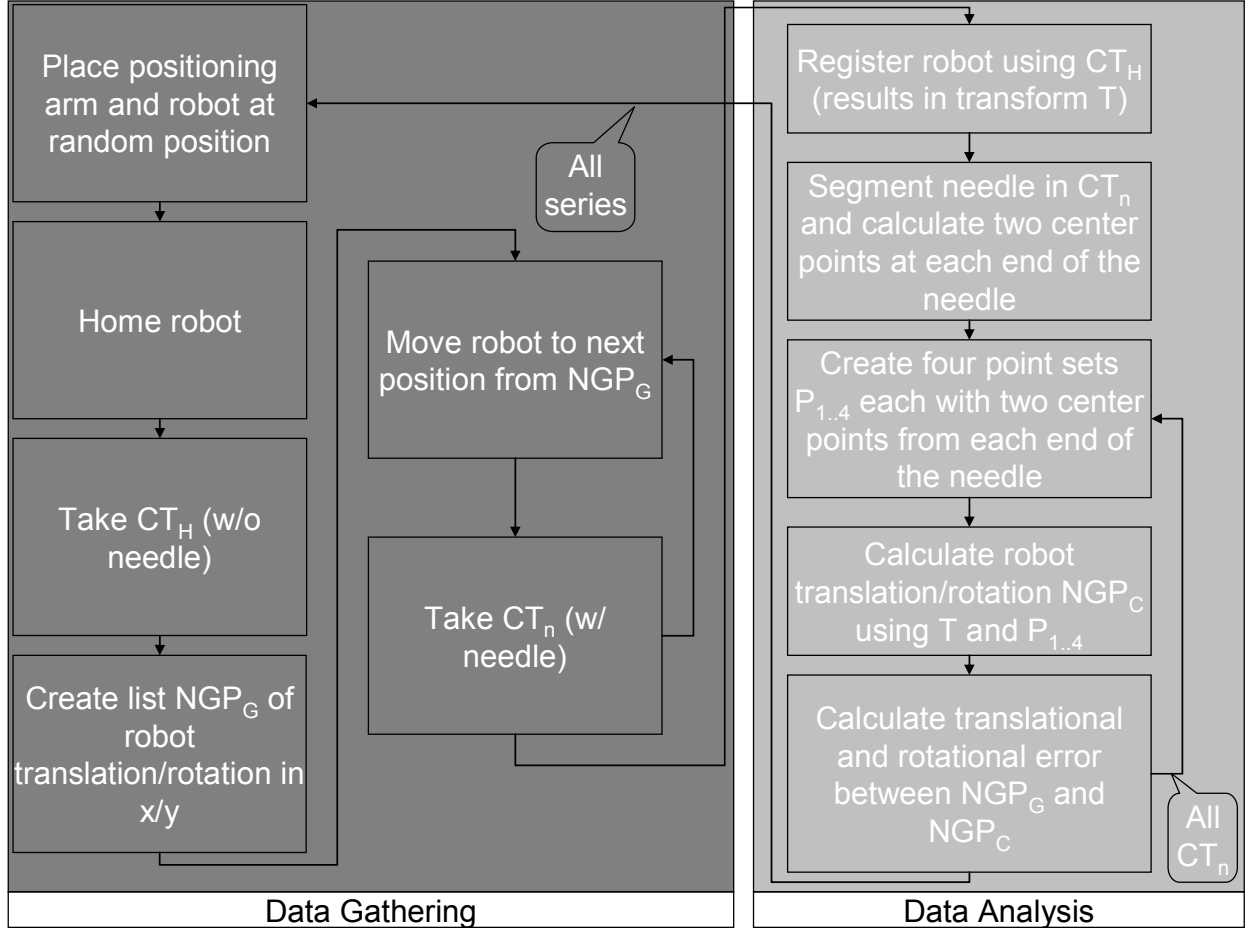
Figure 7. Workflow of experimental setup.

4. RESULTS

Our error calculations of the robot misalignment are shown in Table 1. An estimate of the error between target and instrument tip follows below. The mean translational error over three trials, each with three series, was 0.87 mm. The mean rotational error was 1.46 degrees.

The rotational error was calculated by

$$\alpha = \arctan \sqrt{(\tan \alpha_1)^2 + (\tan \alpha_2)^2} \tag{1}$$

with α_1 and α_2 as the absolute difference between ground truth angle (part of NGP_G) and calculated angle (part of NGP_C) in rotation around x and y, respectively. The rotations around x and y are both centered at the remote center of motion, therefore α is the total rotational error combining α_1 and α_2.

	Series 1	Series 2	Series 3	Average	
Trial 1	0.77	1.16	0.61	0.85	Mean Translational Error (mm)
	1.14	1.60	1.42	1.38	Mean Rotational Error (degree)
Trial 2	1.84	0.74	0.13	0.90	Mean Translational Error (mm)
	1.40	1.34	1.86	1.54	Mean Rotational Error (degree)
Trial 3	0.80	1.05	0.72	0.86	Mean Translational Error (mm)
	1.50	1.22	1.64	1.45	Mean Rotational Error (degree)

Table 1. Mean translational and rotational errors.

Results from trial 2 / series 1 are shown as an example. Four center points, two from each end of the needle, were combined into four point sets $P_{1..4}$ and were used for path planning. Table 2 shows the calculated translation and rotation.

	Translation x (mm)	Translation y (mm)	Rotation x (degree)	Rotation y (degree)
P_1	12.39	10.39	20.06	22.00
P_2	11.86	10.49	20.00	21.66
P_3	11.63	10.38	20.07	21.09
P_4	11.28	10.48	20.01	20.86
Average	11.79	10.44	20.04	21.40

Table 2. Calculated and average translation and rotation for point sets $P_{1..4}$ from trial 2 / series 1.

For this example, the ground truth for translation in x and y was 10 mm and the ground truth for rotation in x and y was 10 degrees. Therefore, the translational error is 1.79 mm in x and 0.44 mm in y; the rotational error is 0.040 degrees in x and 1.40 degrees in y. The combined translational error is 1.84 mm and the rotational error according to (1) is 1.40 degrees.

Although we did not use any phantom or target in our experiments, we also estimate the error between a hypothetical target and the instrument tip by assuming that the rotational and translational error add to each other in the worst case, as follows:

$$e = \sqrt{2d^2 - 2d^2 \cos\alpha} + t = \sqrt{2d^2(1 - \cos\alpha)} + t \tag{2}$$

where d is the instrument insertion depth, α is the rotational error described in (1), t is the translational error, and e is the error between the instrument tip and the target. We used the mean translational error $t = 0.87$ mm and the mean rotational error $\alpha = 1.46$ degrees from above.

The error caused by rotational misplacement is 2.55 mm and the overall error e is 3.42 mm at 10cm insertion depth. Assuming that t and α do not necessarily add to each other during an actual procedure, but rather partly cancel out each other, the mean error between target and instrument tip would be lower.

5. DISCUSSION AND CONCLUSIONS

We have developed an image-guided platform for precision placement of surgical instruments consisting of a custom-designed instrument guide with a built-in fiducial pattern, a portable four degree-of-freedom robot, and custom control software. The platform automatically registers the robot to a CT scan (patient) and positions a surgical instrument for percutaneous access.

To evaluate the algorithmic accuracy, we obtained CT scans from different robot positions and calculated the error based only on these CT scans. The mean error for the overall accuracy evaluation was 0.87 mm for translation and 1.46 degrees for rotation. We also estimated the error between target and instrument tip in a worst case scenario as 3.42 mm at an instrument insertion depth of 10cm. We assume the mean TRE during a phantom experiment would be significantly lower, therefore we will use a phantom in all further experiments.

Therefore, we conclude that our algorithmic approach is suitable for precise instrument placement for minimally invasive interventions, and that the custom instrument guide with a built-in spiral fiducial pattern enables a stable and accurate registration.

Based on discussions with our Interventional Radiology colleagues, this new platform may help physicians hit smaller targets by using image guidance to accurately place surgical instruments. Consequently, the system can make difficult procedures, such as biopsies and radio-frequency ablations, more accurate. In the future, this platform can be modified for any procedure requiring precision placement of instruments, such as gene seed placement.

Furthermore, the Graphical User Interface (GUI), written using the open-source Image Guided Software Toolkit (IGSTK), is very user friendly and has already been used for several tests and demonstrations.

Finally, the robotic hardware system is easy to set up, is extremely portable, and thus would be easily integrated in other environments and adapted to new interventional workflows.

6. ACKNOWLEDGMENTS

REFERENCES

1. M. Marohn: Presentation at Surgery for Engineers 2006. Johns Hopkins University, Baltimore, 2006.
2. B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America A, Vol. 4, p. 629, 1987.
3. K. Cleary, A. Melzer, V. Watson, G. Kronreif and D. Stoianovici. Interventional Robotic Systems: Applications and Technology State-of-the-Art. Minimally Invasive Therapy. 2006; 15:2; 101-113.
4. K. Gary, L. Ibanez, S. Aylward, D. Gobbi, M.B. Blake, K. Cleary. "IGSTK: An Open Source Software Toolkit for Image Guided Surgery." IEEE Computer, 39(4):46-53. April 2006.
5. T. Yoo, D. Metaxas. "Open science--combining open data and open source software: medical image analysis with the Insight Toolkit." Med Image Anal. 2005 Dec; 9(6):506-506.

8.5 Tang 2007: Comparison of Neurosensorimotor …

Reprint begins on the next page and is six pages.

Comparison of Neurosensorimotor Adaptation Under Kinematic and Dynamic Distortions

Jonathan Tang, Jose L. Contreras-Vidal, and Craig Carignan

Abstract— The InMotion2 Robot (Interactive Motion Technologies, Inc.) is used to impart either a kinematic (visual) or dynamic (force) distortion during a series of center-out hand movement tasks. We found that changing distortion type does not differentially affect initial direction error significantly. This suggests that kinematic and dynamic distortion have similar effect on early visuomotor transformations for movement uncorrected by visual feedback. Kinematic distortion affects movement length considerably more than dynamic distortion at only the early stage of learning. We found no statistically significant interactions due to learning from previous exposure. Our tests did show learning during each experiment evident in the time course of metrics. While the eventual goal of this study is to develop robotic interventions for treating patients with neurological disorders, only healthy subjects were used in this study. Future work will expand this study to cover neurologically impaired subjects (Parkinson's Disease and stroke hemiplagia) allowing us to compare the time course of adaptation between impaired and baseline subjects. This data will further help us in our goal to develop large-scale computational models of neuro-sensorimotor adaptation.

I. Introduction

This paper describes an initial pilot study to develop computational models of sensorimotor function of human subjects using robotic platforms. A robotic device is used to impart either a visual (kinematic) or force (dynamic) distortion to a subject's hand as they attempt to conduct a series of center-out tasks on a computer screen while grasping the handle of the robot. Learning rates can then be correlated with physiological data to develop brain models of neurological deficit.

The robot used for this study is an InMotion2 (IM2) Shoulder-Elbow Robotfrom Interactive Motion Technologies, Inc. shown in Figure 1. The IM2 Robot is a four-bar linkage which operates in a planar workspace of approximately 0.5 m^2. The robot uses direct drive motors which can output a continuous maximum force of about 5 lb in each direction depending upon the configuration. The links are constructed of very lightweight tubing which produces low inertial forces as the subject performs a task.

While the eventual goal of this study is to develop robotic interventions for treating patients with neurological disorders such as Parkinson's Disease and stroke hemiplagia, only

Fig. 1. Subjects use the robot handle to move a cursor on the computer screen

healthy subjects were used in this study. This intermediate step is necessary in order to understand, validate, and test the brain model-machine interface. A total of 30 subjects were tested, and data from four were dismissed early on due to incorrect protocol procedure. The subject pool was divided into two groups in which the order of the kinematic and dynamic distortion tests were reversed to study the cross-learning effects.

This article begins with a brief review in Section II of previous work on brain-machine model development using robotic devices. Materials and methods used in this study are described in Section III. The experimental results and analyses are presented in Section IV. Some conclusions are discussed in Section V along with future directions for research.

II. Previous Work

Recent motor control theories suggest that the brain uses internal models to plan and control accurate movements. An internal model is thought to represent how the biomechanics of the arm interacting with the outside world would respond to a motor command; therefore it can be seen as a predictive model of the reafference that helps the system plan ahead [1]. Thus, during adaptation to force fields (external forces applied through the robot manipulandum which alter the normal dynamic characteristics of arm motion), these adaptive internal models are thought to generate compensating torques which allow the arm to track an invariant reference trajectory to a specified target.

There are many different approaches to modeling adaptive sensorimotor behavior ranging from adaptive control techniques to biologically-inspired neural network approaches. A benchmark test performed by many researchers in motor learning is a reaching task between points usually laying along the circumference of a circle at equally spaced intervals (i.e., the center-out task). The human operator is told to move the cursor on the computer screen from point A to point B in a straight line "as fast as possible" using an input device, which is either a computer mouse or robot handle. The experimenter then either distorts the kinematic mapping of the handle or programs the handle to exert environmental force disturbances on the subject. This allows researchers to examine how subjects react to various kinematic and dynamic perturbations thus furthering their understanding of any adaptive processes that might be occurring in parallel

In the case of a distorted kinematic environment (e.g., altered screen cursor-hand relationships), the internal model would represent the new inverse kinematics required to transform a desired movement vector in visuospatial coordinates into a joint-based motor command. Adaptive sensorimotor behavior therefore involves the problems of trajectory planning, coordinate transformation, and control, and the brain must solve these problems to bring the hand from the starting position to a desired target location [2]. In the case of dynamic distortions, the robot handle exerts forces back on the subject as s/he attempts to move between points. This causes the subject to compensate for the disturbance by learning the external forces and modulating force output to the handle.

An interesting result of these studies has been that despite large differences in models, they often display three common features: a trajectory generator, sensory feedback and control loops, and an adaptive process. Models of reach planning based on either kinematic (e.g., velocity command model) or dynamic (e.g., torque command model) variables have been proposed in the literature. In the velocity command model, the motor command is specified as joint angle velocities, whereas in the torque command model, the trajectory generator outputs a set of nominal joint torques for the arm which are learned over a lifetime of executing such tasks [3]. The second feature is that humans will use visual and kinesthetic feedback to correct for the arm motion when it drifts from this nominal path. Finally, humans will adapt to environmental disturbances or distortions to the sensory mappings to keep the arm moving along the desired trajectory [4][5].

In previous work, we have used virtual environments to assess how the brain learns internal models of changing environments [6][7][5][8]. In these studies we used kinematic distortions (e.g., rotations of the screen cursor motion representing hand movement) to show that young healthy subjects performing center-out movements are able to adapt their internal model of the visuomotor transformation for aiming in response to exposure to a 90 deg screen cursor rotation. Specifically, during early-exposure trials, subjects perform spiral-like movements using indirect visual feed-back of movement (using the screen cursor movement); however, with practice, these movements become straighter and smoother by late-exposure. Interestingly, during post-exposure, after removal of the rotational distortion, subjects show spiral-like movements in the opposite direction as in the early-exposure. These after-effects are thought to be indicative of adaptation of the internal visuomotor transformation during exposure trials.

III. Materials and Methods

A. Apparatus

The robot used for this study (Figure 1) is an InMotion2 (IM2) Shoulder-Elbow Robot from Interactive Motion Technologies, Inc. The IM2 robot is FDA-approved as both an evaluation and therapy device. It has direct-drive motors at the base which transmit forces to the handle through a four-bar linkage as shown in Figure 1. The handle moves within a rectangle of dimension 90 x 60 cm just above the surface of the table. The maximum continuous force output of the robot at the handle is approximately 30 N in each direction. The handle is pinned to the end of the last link providing a third, unactuated degree of freedom. The robot control computer runs a real-time linux operating system to read handle position, apply force to the handle, draw cursor and targets, and calculate kinematic and dynamic transformations.

B. Subjects

30 neurologically unimpaired subjects (18 - 50 years old) were recruited from Georgetown University (GU) and the University of Maryland (UMD). All subjects gave written informed consent in accordance with the Internal Review Boards of GU, UMD, and our sponsoring agency US Army Telemedicine and Advanced Technologies Research Center (TATRC).

C. Procedure

Subjects faced a flat computer display (41 x 30cm) situated in front of them at a distance of 60 cm and at eye level (as in Figure 2). Vision of the arm/hand/robot was be prevented with a white board (not shown). Subjects performed horizontal point-to-point movements with their right, dominant arm (the so-called center-out task). The subjects controlled the movement of the screen cursor by moving their arm in the horizontal plane. Subjects were instructed to make point-to-point movements as fast and straight as possible by moving the screen cursor from a central starting location (presented as a circle of 10mm diameter) to one of four yellow target circles (diameter, 10mm; directions, $45°$, $135°$, $225°$, and $315°$) chosen at random and displayed on the screen. Time to target presentation was random (600ms to 1000ms) and if the subject had not hit the target within two seconds after leaving the center circle, the target turned red indicating that the subject should speed up.

Each experimental session consisted of 260 trials split into four phases, 1) pre-exposure (no recording), 2) pre-exposure,

3) exposure, and 4) post-exposure as shown in Table III-C. 20 practice trials were administered at the beginning of the session to familiarize the subjects with the setup. Trials 21-60 (pre-exposure phase) were performed under no visual or force pertubations. Trials 61-240 (exposure phase) were performed under kinematic or force distortion. Trials 241-260 (post-exposure phase) were performed under normal conditions to test for after-effects. Subjects participated in both kinematic and dynamic distortion experiments, but were randomly assigned to do either the kinematic or dynamic distortion first.

TABLE I
EXPERIMENTAL PROTOCOL

Trial #	Kinematic	Dynamic
1-20	not recorded	not recorded
21-60	pre-exposure	pre-exposure
61-240	kinematic distortion	dynamic distortion
241-260	post-exposure	post-exposure

1) Kinematic Distortion Trials: Under kinematic distortion, the transformation between cursor and hand movement was rotated by 60 degrees clockwise as shown in Equation 1 and Figure 2. Vision of the arm/hand/robot was prevented with a white board (not shown).

$$\begin{bmatrix} X_c \\ Y_c \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X_h \\ Y_h \end{bmatrix} \qquad (1)$$
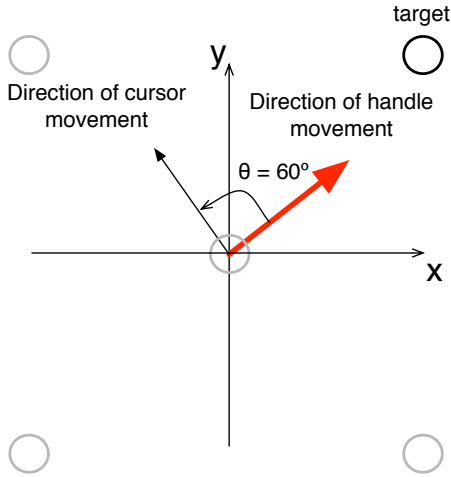

Fig. 2. A counterclockwise rotation of $60°$ was applied between the direction of the cursor and robot handle in the kinematic distortion trials.

2) Dynamic Distortion Trials: Subjects performed the same center-out task as in kinematic distortion experiments except a curl field (Shadmehr and Mussa-Ivaldi, 1994) was applied to the subject's hand by the robot handle and no white board is used to block vision. Specifically, the hand was subjected (during exposure trials) to the field F, as shown in Equation 2 and Figure 3, where $\{F_x, F_y\}$ represents the force vector applied to the robot handle, $\theta = 60°$ is the rotation angle, $\{V_x, V_y\}$ is the velocity vector of the handle, and k=13 N/m is the stiffness coefficient.

$$\begin{bmatrix} F_x \\ F_y \end{bmatrix} = k \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} V_x \\ V_y \end{bmatrix} \qquad (2)$$
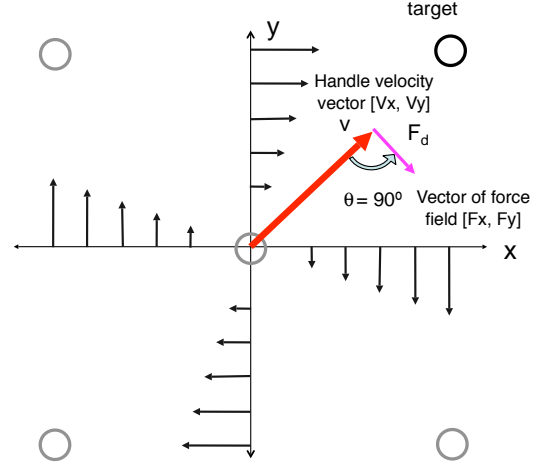

Fig. 3. A counterclockwise curve field proportional to the velocity of the handle was applied in the dynamic distortion trials.

IV. RESULTS AND ANALYSIS

A. Stages of Learning

Subjects perform a total of 260 trials that can be analyzed at the following areas of interest shown in figure 4 and described below:

1) Familiarization: 20 unrecorded movements to become familiar with the system.
2) Baseline: 40 movements during pre-exposure with recording turned on to establish baseline.
3) Early exposure: 20 movements made right after enabling kinematic or dynamic distortion and the neuro-sensorimotor model is adapting to distortion.
4) Exposure: 140 movements made during intermediate learning.
5) Late exposure: 20 movements made during late learning at the end of the exposure phase.
6) After effects: 20 movements made post exposure when distortion is turned off and the subject's neuro-sensorimotor model is re-adapting to normal conditions.

Figures 4 and 5 show an example of the movements taken by one subject during kinematic and dynamic distortion. We can see the increase in trajectory error in early exposure, the improvement due to adaptation to distortion in late exposure, and another increase in trajectory error in post exposure due to after effects of distortion. The three stages that we have selected for further analysis are 1) early exposure, 2) late exposure, and 3) after effects because they represent the periods of early, late adaptation, and re-adaptation respectively.
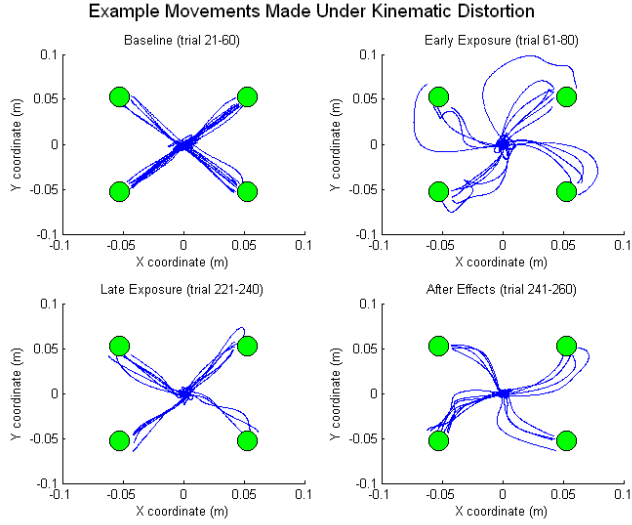
Fig. 4. Movements made by one subject under kinematic distortion. Focus is on baseline, early exposure, late exposure, and aftereffects
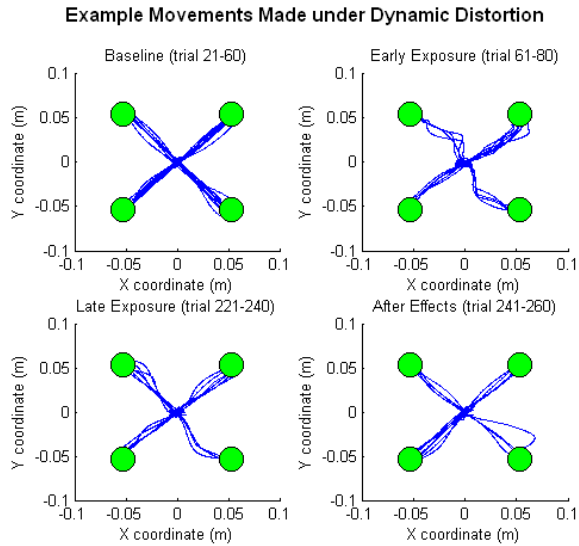

Fig. 5. Movements made by one subject under dynamic distortion. Focus is on baseline, early exposure, late exposure, and aftereffects

B. Metrics

Handle position is sampled at 200 hertz and filtered by a 2nd-order low-pass Butterworth filter at 5 hz. Several metrics were calculated using the position data including, 1) initial direction error, 2) path length, 3) movement time, 4) maximum velocity. Movements returning to the center were not analyzed. The initial direction error (IDE) was calculated as the angle between the initial direction of planned trajectory and the direction to the target. A vector from the center to the position of the handle at 2 cm after movement onset determined the initial direction of planned trajectory. The IDE is thought to represent early visuomotor transformations

for movement uncorrected by visual feedback [8]. The path length calculated as the total distance traversed to reach the target, which is a measure of visuomotor adaptation including both uncorrected and corrected movement.

One of our goals is to investigate and compare the time course of adaptation to kinematic (visual) versus dynamic (force) distortions. It seems the best way to accomplish these would be to standardize the scores (e.g., IDE, path length, velocity, etc). For each of the metrics, we took the average and standard deviation of the baseline trials (21- 60), and used the corresponding mean and STD to standardize the rest of the trials. For example for IDE:

$$\text{NormalizedIDE(trial)} = \frac{\text{IDE(trial)} - \text{mean(baseline)}}{\text{std(baseline)}} \tag{3}$$

Now the units for normalized metrics would be in +/- standard deviations from the baseline mean. Using these normalized metrics, we can now aggregate all subjects, and compare between kinematic and dynamic groups.

C. Curve Fitting

To measure the time course of adaptation, we applied a double exponential fit to the exposure trials (60-240),

$$Ae^{\frac{-t}{B}} + Ce^{\frac{-t}{D}} \tag{4}$$

where, A is the magnitude of trajectory error and B is the learning rate in the early exposure phase, and C and D are the corresponding coefficients in the mid to late exposure phase. A single exponential is used to fit the post-exposure trials (241-260) due to the smaller amount of trials.

To minimize the influence of outliers, we used robust least squares regression with bisquare weights. Robust fitting uses an iteratively reweighted least squares algorithm, which minimizes a weighted sum of squares, where the weight given to each data point depends on how far the point is from the fitted line.

Figure 6 and 7 show the aggregate charts for initial direction error and path length. Note the 3 areas of interest in the charts for IDE and path length as described previously in the beginning of Section IV. We see marked increase in IDE and path length at the onset of distortion, with gradual learning (decaying at exponential rates) throughout the rest of the experiment, and finally another increase in IDE and path length when distortion is removed.

D. Statistics

The total population consists of two general groups.

1) Kinematic distortion first - subjects who were tested with kinematic distortion first, and dynamic distortion second.
2) Dynamic distortion first - subjects who were tested with dynamic distortion first, and kinematic distortion second.

Of these 30 participants, data from 4 subjects were discarded due to non-compliance to protocol, 16 subjects were
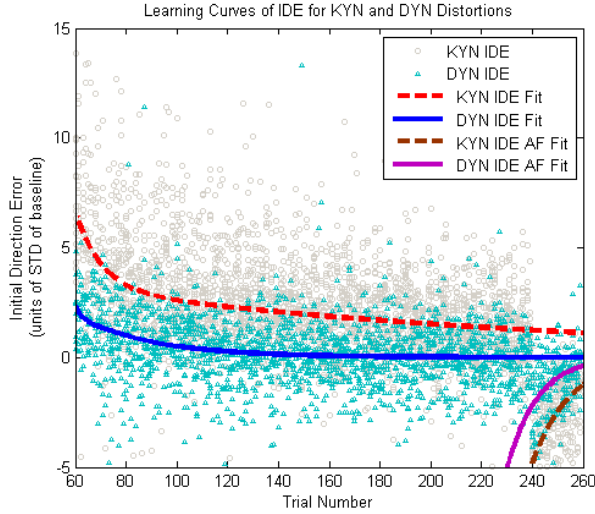
Fig. 6. Curve fits of the median initial direction error from kinematic and dynamic data, using robust double exponentials fit on exposure trials, and single exponential on post exposure.
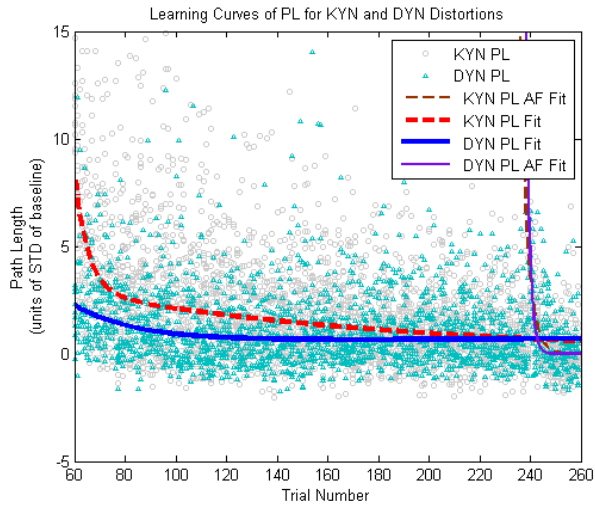

Fig. 7. Curve fits of the median path length from kinematic and dynamic data, using robust double exponentials fit on exposure trials, and single exponential on post exposure.

in the kinematic distortion first group, and 10 subjects were in the dynamic distortion first group. The data can therefore be split into four groups as follows:

1) A - (16 sets) kinematic data from kinematic first group.
2) B - (10 sets) dynamic data from kinematic first group.
3) C - (10 sets) dynamic data from dynamic first group.
4) D - (16 sets) kinematic data from dynamic first group.

The first goal is to compare the adaptation to kinematic versus dynamic distortions at three stages (early, late, aftereffects) of learning. Using a 2-way general linear model repeated measure ANOVA on the four groups (A, B, C, D) defined previously, we can determine the significance of distortion type and stage of learning on IDE and path

length metrics. We find that distortion type does not affect IDE significantly (p = 0.15) at all three stages of learning. The Bonferroni corrected multi-comparison results show that there's significant difference at the different phases (early, late, after) but no signficant effect from distortion types. However distortion effects are significant for path length metrics. (p = 0.0024) The Bonferroni corrected multi-comparison results show that kinematic distortion affects path length significantly more than dynamic distortion at only the early stage of learning (early exposure). Box plots in Figure 8 and 9 illustrate the mean, median, and quantile range for IDE and path length metrics in groups A and C.

A second goal is to examine if adaptation to one distortion would affect the other in a sequential manner. In the second test, we compared kinematic data from the kinematic first and dynamic first groups (A and D) using an ANOVA and found no statistically significant interactions due to learning from previous exposure.

The third analysis was a one-sample t test for aftereffects, showing that all groups learnt the task and exhibited significant aftereffects (p < 0.007).
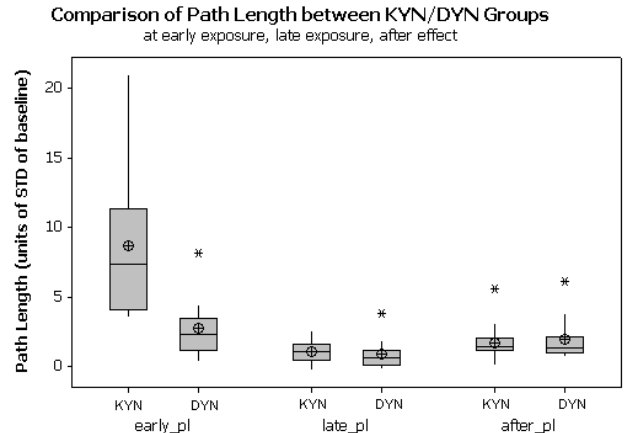
Fig. 8. Box plot comparisons of path length at early, late, and after effect stages under kinematic or dynamic distortions

V. CONCLUSION AND FUTURE WORK

This paper describes an initial pilot study to develop computational models of sensorimotor function of human subjects using robotic platforms. The robot used for this study is an InMotion2 (IM2) Shoulder-Elbow Robot from Interactive Motion Technologies, Inc. The robotic device is used to impart either a kinematic or dynamic distortion to a subject's hand during a series of center-out tasks. While the eventual goal of this study is to develop robotic interventions for treating patients with neurological disorders such as Parkinson's Disease and stroke hemiplagia, only healthy subjects were used in this study. This intermediate step is necessary in order to understand, validate, and test the brain model-machine interface.

Our first goal was to to investigate the time course of adaptation to kinematic versus dynamic distortions. To do

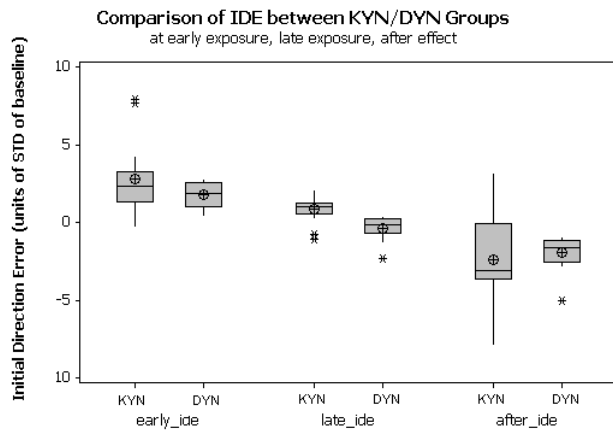**Comparison of IDE between KYN/DYN Groups**
at early exposure, late exposure, after effect

Fig. 9. Box plot comparisons of initial direction error at early, late, and after effect stages under kinematic or dynamic distortions

REFERENCES

[1] H. Imamizu and S. Miyachi, "Human cerebellar activity reflecting an acquired internal model of a new model," *Nature*, vol. 403, pp. 192–195, 2000.

[2] H. Imamizu, Y. Uno, and M. Kawato, "Adaptive internal model of intrinsic kinematics involved in learning an aiming task." *J Exp Psychol Hum Percept Perform*, vol. 24, no. 3, pp. 812–29, 1998.

[3] S. Sober and P. Sabes, "Multisensory integration during motor planning," *J Neurosci*, vol. 23, no. 18, pp. 6982–92, 2003.

[4] R. Shadmehr and F. Mussa-Ivaldi, "Adaptive representation of dynamics during learning of a motor task," *J Neuroscience*, vol. 14, pp. 3208–3224, 1994.

[5] J. Contreras-Vidal and E. Buch, "Effects of parkinson's disease on visuo-motor adaptation," *Exp Brain Res*, vol. 150, pp. 25–32, 2003.

[6] F. Kagerer, J. Contreras-Vidal, and G. Stelmach, "Adaptation to gradual versus sudden visuo-motor perturbations," *Experimental Brain Research*, vol. 115, pp. 557–561, 1997.

[7] J. Contreras-Vidal and S. Kerick, "Independent component analysis of dynamic brain responses during visuomotor adaptation," *Neuroimage*, 2004.

[8] E. Buch, S. Young, and J. Contreras-Vidal, "Visuomotor adaptation in normal aging," *Learning and Memory*, vol. 10, pp. 55–63, 2003.

this we fitted learning curves to initial direction error and path length metrics and determined whether there were statistically significant interactions at early, late, and after effect stages of learning. We found that distortion type does not affect IDE significantly at all three stages of learning. This lead us to believe kinematic and dynamic distortion have similar effect on initial direction error, which are indicative of early visuomotor transformations for movement uncorrected by visual feedback. However distortion effects were significant for path length metrics. The Bonferroni corrected multi-comparison results show that kinematic distortion affects path length significantly more than dynamic distortion at only the early stage of learning (early exposure). This seems to indicate that kinematic distortions significantly affect the internal model that uses visual feedback for trajectory correction only at the early stage of learning.

Our second goal was to examine the sequential effects of adaptation, whether an earlier session would affect the other in a sequential paradigm. We found no statistically significant interactions due to learning from previous exposure. Our tests did show significant learning during each experiment, as seen from the time course in path length and IDE metrics between stages of learning.

Future work will expand this study to cover neurologically impaired subjects (Parkinson's Disease and stroke hemiplagia) allowing us to compare the time course of adaptation between impaired and baseline subjects. This data will further help us in our goal to develop large-scale computational models of neuro-sensorimotor adaptation. Ultimately this research will help us develop robotic interventions for treating patients with neurological disorders.

8.6 Wilson 2007: Hardware and Software Assessment …

Reprint begins on the next page and is eleven pages.

A hardware and software protocol for the evaluation of electromagnetic tracker accuracy in the clinical environment: a multi-center study

Emmanuel Wilson[*a], Ziv Yaniv[a], Hui Zhang[b], Christopher Nafis[c], Eric Shen[d], Guy Shechter[d],
Andrew D. Wiles[e], Terry Peters[e], David Lindisch[a], Kevin Cleary[a]

[a]Imaging Science and Information Systems (ISIS) Center, Department of Radiology, Georgetown University Medical Center, 2115 Wisconsin Avenue, Washington, DC 20007, USA
[b]Accuray, 1310 Chesapeake Terrace, Sunnyvale, CA 94089, USA
[c]GE Research, 1 Research Circle, Niskayuna, NY 12309, USA
[d]Philips Research North America, 345 Scarborough Rd., Briarcliff Manor, NY 10510, USA
[e]Imaging Research Laboratories, Robarts Research Institute and Dept. of Medical Biophysics, The University of Western Ontario, 100 Perth Drive, London, Ontario, N6A 5K8, Canada

Abstract

This paper proposes an assessment protocol that incorporates both hardware and analysis methods for evaluation of electromagnetic tracker accuracy in different clinical environments. The susceptibility of electromagnetic tracker measurement accuracy is both highly dependent on nearby ferromagnetic interference sources and non-isotropic. These inherent limitations combined with the various hardware components and assessment techniques used within different studies makes the direct comparison of measurement accuracy between studies difficult. This paper presents a multi-center study to evaluate electromagnetic devices in different clinical environments using a common hardware phantom and assessment techniques so that results are directly comparable. Measurement accuracy has been shown to be in the range of 0.79-6.67mm within a 180mm^3 sub-volume of the Aurora measurement space in five different clinical environments.

Keywords: electromagnetic tracking, accuracy analysis, evaluation protocol, clinical environment

1. Introduction

Spatial measurement systems play an integral role in many image-guided surgery (IGS) systems. Of the several types of spatial measurement systems, optical tracking systems were the first to make inroads into clinical practice[1,2]. Over the years, as computation capabilities increased, research and deployment of IGS systems has expanded across many medical disciplines. From a clinical standpoint, this change is attributable to the proliferation of minimally-invasive surgery techniques. As incisions become smaller, the use of flexible surgical tools such as guidewires, catheters and small-bore needles has become more prevalent in standard practice. Optical trackers have two key requirements that limit their use: reliance on direct line-of-sight between the optical tracking system and tracked tool and the ability to only track rigid tools. For these reasons, interest in electromagnetic tracking as a technology that can potentially overcome these limitations has begun to grow[3,4].

For all their benefits, electromagnetic tracking systems have not yet found the same widespread acceptance as optical systems. To a large extent this is because of the susceptibility of electromagnetic tracker measurements to distortion by metal and electromagnetic interference sources[5,6]. The quantification of these errors has been a focus of several groups developing IGS systems. Direct comparison of results obtained by different groups is difficult because of the differing approaches taken to devise a hardware phantom and methods to analyze and quantify the data[7-12]. This paper presents initial work towards defining both a hardware phantom and assessment methods to quantify the accuracy of the Aurora

---

* wilsone@isis.georgetown.edu

electromagnetic tracker (Northern Digital Inc., Waterloo, Ontario, Canada). We present results from four institutes. The work is seen as ongoing, with the main intention being to present initial work at defining a standard phantom and method of analysis for the benefit of a wider audience.

At our institute (Georgetown University Medical Center), we have gained extensive experience in the targeting of liver tumors for needle biopsy in swine models using electromagnetically navigated guidance[13-15]. Based on our experience through this work, a set of primary requirements were devised for the design of the hardware phantom and are enumerated below:

1. The phantom must be affordable and easily replicated.
2. The phantom must account for the entire volume of interest, yet not be unwieldy.
3. Electromagnetic device and phantom setup must be simple and efficient.
4. Data collection time should be kept to a minimum.

Based on the body of literature available on accuracy analysis of electromagnetic devices and our experience, a set of requirements for the analysis of the data and quantification of errors were devised and are listed below:

1. Statistics used to define the accuracy must be representative of the underlying nature of electromagnetic measurement errors.
2. The result must be concise yet representative of errors within a volume.
3. Software data collection and analysis tools must be cross platform (portable across operating systems).

Given these requirements, we designed the hardware phantom and analysis protocol described in the following section.

2. Materials and Methods

2.1 Measurement device

This multi-center study was conducted solely on the NDI Aurora electromagnetic tracking system (NDI Inc, Waterloo, Canada). The device consists of a flat field generator, a system control unit that interfaces with a PC, and tracked sensor coils. The device computes the position and orientation of the sensor coils at a maximal rate of 40 Hz based on the relative strength of pulses from the field generator induced within the sensor coil. A single five degree-of-freedom MagTrax sensor needles (Traxtal Technologies, Bellaire, TX, USA) was used for the study. The NDI Aurora device has a manufacturer quoted RMS positional accuracy between 0.7 - 0.8mm and RMS orientation accuracy of 0.3° within a $200 - 400$mm radial distance from the field generator[16].

2.2 Hardware phantom

The hardware phantom used for this study is a plexiglass cube of 180mm sides with 225 precisely machined holes from the top face, spaced uniformly apart with 10mm spacing. The holes range in depth from 10mm to 150mm in random order. All holes were machined to a diameter of 1.4mm so as to provide a snug channel for a 5DOF MagTrax needle with an outer diameter of 1.27mm. This design addresses all the requirements enumerated above.

The simple cube design results in reduced machining costs and is easily replicated at other locations as it requires only basic machining capabilities to machine the holes.

The phantom's spatial extent, 180mm^3, encompasses a sufficient volumetric region for most interventional procedures. This is based on our observations of clinical needle biopsy and tumor ablation procedures conducted within the thoracic-abdominal cavity of human patients. At the same time this size is sufficiently small as to not make the phantom unwieldy.

The dimensions of the cube make it portable and easy to setup within the clinical environment, with no additional hardware components required for phantom placement.

Finally, the choice of 225 spatial samples minimizes the data collection time to approximately 30-45 minutes while still providing a consistent description of the underlying errors. This spatial distribution and sample size were deemed sufficient based on our previous work described in[11]. In that work we evaluated the dependence of various statistics on the number of sampling points uniformly distributed in the volumetric region of interest. The dataset consisted of 2548 data nodes collected within a 240mm³ volume. Various sample sizes were selected from the larger dataset and the same analysis methods proposed here were used on each subset. The statistic measure for each subset, summarized in Table 1, provides a quantitative outlook on how each statistic varies with the number of sample points. Based upon these results, it was decided that 225 nodes are sufficient to characterize the underlying error distribution, while providing feasible data collection time.

Number of Nodes	2500	1000	500	250	225	100	50
RMS Error	2.38	2.16	2.42	2.38	2.39	2.08	1.94
Median Error	1.79	1.59	1.79	1.79	1.86	1.63	1.26
Standard Deviation	1.24	1.12	1.33	1.27	1.24	1.09	1.13
Max Error	6.97	6.00	6.96	6.08	5.97	4.59	5.09
Error Range	6.93	5.88	6.91	6.00	5.93	4.56	4.93

Table 1: Summary of effect on statistic measures based on number of nodes selected.

Due to unforeseen mechanical limitations of machining 1.4mm diameter holes to depths larger than 30mm, it was decided to construct the cube from 36 layers of cast acrylic of 3/16 inch thickness, each stacked upon one another to form the cube. Being cast acrylic, the sheets have a width that is variable in the range of + 0.8mm to – 1.5mm. The holes were machined using a EuroLaser M-1200 laser cutting system (eurolaser GmbH, Seevetal, Germany) to a hole diameter accuracy of 0.0025mm. Accuracy of hole position along the X-Y plane is assumed to be the same as the positioning device used during laser cutting with an accuracy of 0.005mm. The depth of each hole is variable because of the large width tolerance of the cast acrylic sheets. Therefore, a Mitutoyo mechanical depth gauge with a resolution of 0.01mm and instrumentation error under 0.03mm (Mitutoyo Corporation, Japan) was used to record the depth of all holes. Using the recorded depth, the end point of each hole was assigned a Cartesian coordinate based on an arbitrarily chosen axes with origin at the vertex closest to hole number 1 (Figure 1). Holes are numbered 1 to 225 and it is assumed that data collection is performed in sequential order as registration between the measured block coordinates and Aurora data is done on similar datasets.
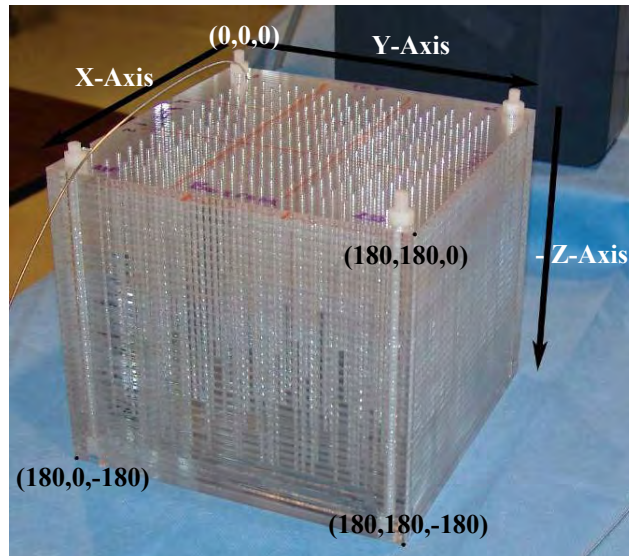

Figure 1: Hardware phantom and its relative coordinate.

2.3   Analysis method

Various error statistics have been used in previous evaluation studies to express the measurement accuracy of the electromagnetic tracking devices. The goal of our analysis method was to provide a concise set of statistics that describe the errors within the volume of interest. In addition to ferromagnetic distortion sources, accuracy of the Aurora device is further hampered by separation between the field generator and sensor coils, such that position errors are non-isotropic and follow a lognormal distribution. This has been established in previous work and position errors have been shown to most closely represent a lognormal distribution[6,11,17]. Therefore, measures such as mean and standard deviation alone are not adequate representations of the error. For this reason, the geometric mean and geometric standard deviation were chosen along with the median, maximal error and RMS error measures. The geometric standard deviation and geometric mean together define the lognormal probability density function in the same way as mean and standard deviation do for a Gaussian distribution. Formulas for all of the statistics used in our analysis are given in Table 2.

(a)
$$E_m = \sum_{i=1}^{n} \varepsilon_i$$

(d)
$$\Pi = \sqrt[n]{\varepsilon_1 \bullet \varepsilon_2 \bullet \varepsilon_3 \bullet ... \bullet \varepsilon_n}$$

(b)
$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\varepsilon_i - E_m)^2}$$

(e)
$$\sigma_g = \exp\left(\sqrt{\frac{\sum_{i=1}^{n} (\ln \varepsilon_i - \ln \Pi)^2}{n}}\right)$$

(c)
$$E_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^2}$$

(f)
$$p(\varepsilon) = \frac{1}{\varepsilon \ln \sigma_g \sqrt{2\pi}} e^{\frac{-(\ln \varepsilon - \ln \Pi)^2}{2(\ln \sigma_g)^2}}$$

Table 2: Statistics used in data analysis, (a) mean, (b) standard deviation, (c) RMS, (d) geometric mean, (e) geometric standard deviation, (f) lognormal probability density function.

To minimize the effect of measurement noise, at each collection node, 100 data samples are logged. Using the standard deviation between samples as a marker, 100 samples at each node were found to be sufficient. Additional samples did not provide a substantial improvement in spatial measurement to counter the lengthened collection time, as shown in Table 3.

	No. of Data Samples (at each node)						
	10	50	100	150	200	250	300
Mean StdDev	0.186	0.405	**0.412**	0.410	0.409	0.410	0.409
Min StdDev	0.013	0.015	**0.014**	0.013	0.012	0.011	0.011
Max StdDev	1.413	2.647	**2.641**	2.554	2.495	2.487	2.466

Table 3: Summary of variation in standard deviation based on choice of number of samples.

Absolute position error is evaluated using Horn's point based registration algorithm[18]. Transforming the measured Aurora readings to the arbitrary phantom coordinates and computing the difference between the two provide a representation of position error at each node. Registration nodes are selected in a pseudo-arbitrary manner. A set of eight to ten registration nodes that fall within the "optimal" measurement sub-volume of the Aurora device are selected. It is assumed that these points suffer from minimal distortion and can be used for registration purposes (this is validated during the analysis). Note that the "optimal" location will vary depending on device setup.

In addition to these statistics, a graph depicting the position errors as a function of distance separation between the field generator and sensor and a histogram plot of position errors at all nodes are provided for each test, examples of which

can be seen for the clinical suite results in Section 3.2. These plots provide visual confirmation of the nature and distribution of the errors across the volume.

2.4 Data collection and setup method

A software data-logging utility that is portable between different operating system environments was developed using the open-source IGSTK framework [19]. The goal in the longer term is to make the software code and hardware design freely available for download so that other groups wishing to test the accuracy of electromagnetic devices can do so using the protocol set forth here using the same data collection methods.

The first part of the setup process is placement of the phantom such that the occupied volume coincides with a clinical volume of interest. In the following we give an example setup to assess electromagnetic tracker accuracy in a region of space within our Interventional Radiology suite that coincides with the usual location of the liver when a patient lies supine on the couch. This is followed by placing the field generator, affixed onto the Aurora mounting arm, at a location that remains clinically viable. In our experience, in a standard Interventional Radiology suite, we have found a separation of 330mm between front face of the field generator and center of the phantom and field generator at a height of 280mm above the Interventional Radiology patient table to be optimal in terms of minimal distortion from the metal components of the patient table while remaining clinically feasible for liver biopsy procedures. A typical setup of the field generator and phantom within a DynaCT environment at Georgetown University Hospital is shown in Figure 2, where (a) denotes the 280mm distance between the field generator and the table and (b) denotes the distance between field generator face and center of the phantom. The exact positioning of the phantom and field generator will be dependent on the intended clinical application and these distances are intended only as a suggestion. Field distortion from nearby metal components can be greatly different in another clinical environment and it is quite likely that alternate setup positions exist even in our clinical environment that are both clinically feasible and possibly result in reduced distortion artifacts.

Once the phantom and field generator are satisfactorily placed, the MagTrax needle is fully inserted into each hole in the numbered order. The data-logger collects 100 samples at each node and outputs it to individual text files. One data collection pass is completed in approximately 30-45 minutes.

Figure 2: Typical device placement for tests in DynaCT suite at Georgetown University Medical Center.

3. Results

Electromagnetic tracker accuracy evaluation was carried out in two types of environments: 1) baseline tests in a ferromagnetically clean environment and 2) clinical environments. The clinical environment varied between groups and as expected, so did the accuracy of the device due to varying types and amount of unavoidable metallic distortion components specific to each environment.

3.1 Baseline results

Results from tests done in a ferromagnetically clean laboratory environment are summarized in Table 4 (all distance and error measures are provided in millimeters). As it has been established through previous work that electromagnetic tracker accuracy is highly dependent on the separation between the field generator and sensor, the distance between the field generator and center of the phantom is presented within the second column of the table as an indicator of the severity of the measurement errors. The larger errors within the Georgetown environment are due to the tests being conducted in a laboratory with several computers, monitors, fluorescent lighting and other possible sources of interference. While every attempt was made to eliminate metal objects within a one meter radial distance from the field generator, the larger errors within this environment are indicative of unavoidable distortion sources.

Test location	FG-phantom separation	RMS error	Median error	Standard deviation	Max error	Range	Geometric mean	Geometric std dev
Georgetown University - lab	190	1.54	1.06	0.9	6.67	6.5	1.06	1.75
Philips Research	190	0.74	0.49	0.44	4.17	4.12	0.49	1.87
GE Research	150	0.38	0.26	0.23	1.97	1.92	0.25	1.82
GE Research	285	0.32	0.24	0.17	1.8	1.78	0.23	1.8
GE Research	440	1.27	1.02	0.63	3.42	3.33	0.92	1.9

Table 4: Summary of baseline tests conducted at three of the four institutes involved in this study. Second column provides the distance between the Field Generator (FG) and the face of the phantom. Baseline tests were not conducted at one site (CSTAR). All measurements are in millimeters.

3.2 Clinical suite results

A greater variability in the measurement accuracy of the electromagnetic device is visible within the clinical environment. Generally speaking, the errors were consistently greater within the clinical environment, much as would be expected, except at Georgetown University. The relatively high accuracy in the DynaCT suite at Georgetown University may be attributable to the fact that there are minimal metal artifacts in the vicinity of the measurement volume. The test location is the patient table which is composed of carbon fiber and is suspended further away from the electronic components that drive the patient table stage.

The entire set of results is summarized in Table 5, where all distance and error measures are provided in millimeters. The error-vs-distance and histogram plots and picture of experiment setup of one test from each group are presented in Figures 3 and 4. While it is hard to draw succinct conclusions about the source of measurement errors, these results provide a quantitative outline of the errors under the specific test conditions. In addition, the results also provide insight into the relationship between errors and the distance between field generator and sensor, from which one can better assess what the limitations are for any clinical application that may rely on electromagnetic tracker data within that particular environment.

Test location	FG-phantom separation	RMS error	Median error	Standard deviation	Max Error	Range	Geometric mean	Geometric std dev
IR Suite – DynaCT (Georgetown)	245	0.79	0.53	0.43	2.71	2.67	0.54	1.91
Bronchoscopy Suite (Georgetown)	270	6.67	4.98	3.57	20.86	20.31	4.65	1.9
CT Suite (Philips)	205	4.19	2.82	2.53	22.24	21.95	2.74	1.87
OR Suite – no lights (CSTAR Animal OR)	195	3.14	2.39	1.46	9.5	9.37	2.45	1.7
OR Suite – w/ lights (CSTAR Animal OR)	195	3.07	2.35	1.46	9.45	9.31	2.37	1.71
OR Suite – w/ lights and surgical tools (CSTAR Animal OR)	195	2.92	2.29	1.3	8.2	8.01	2.33	1.65
C-Arm (GE)	285	1.15	0.84	0.55	3.19	3.09	0.87	1.75

Table 5: Summary of clinical accuracy evaluation tests conducted at various institutes within different environments. Second column provides the distance between the Field Generator (FG) and the face of the phantom. All measurements are in millimeters.
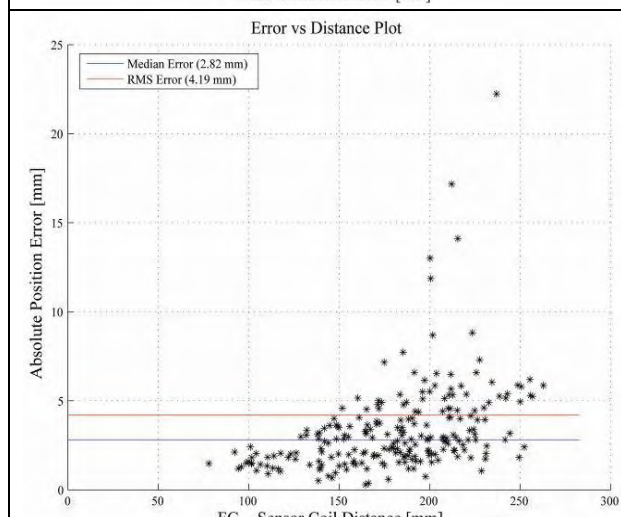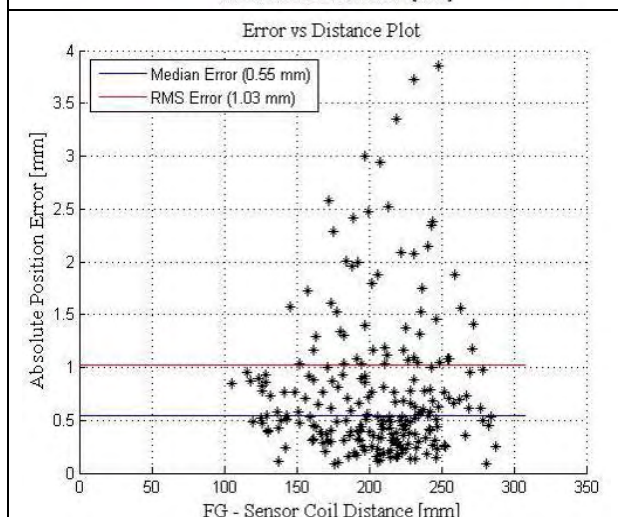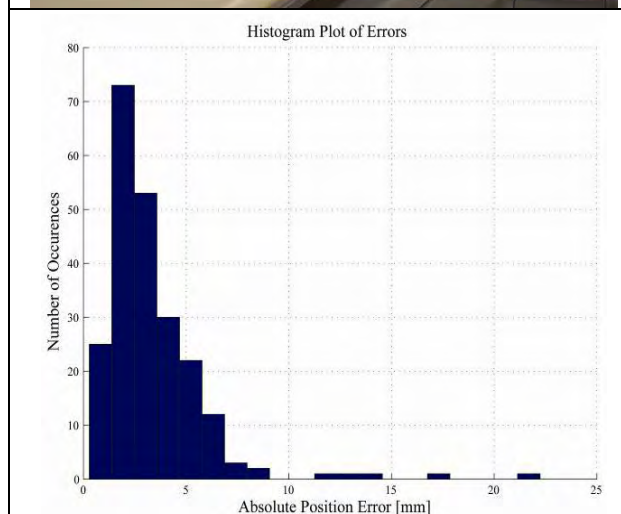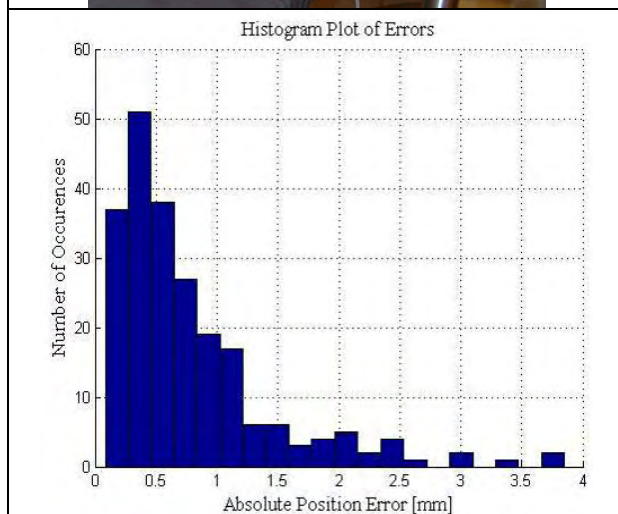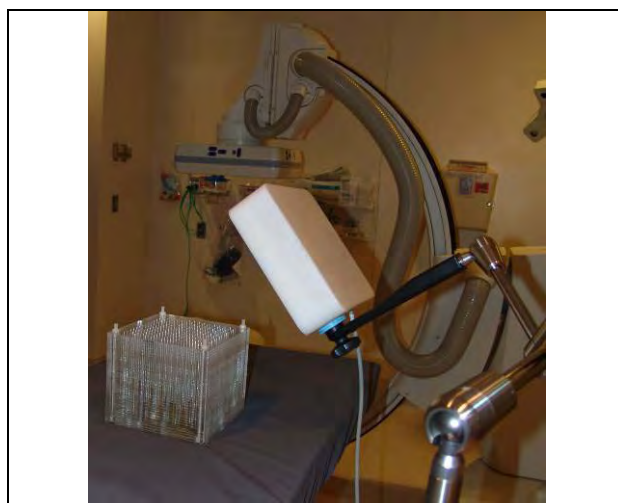
Figure 3: Summary of results (a) IR suite – DynaCT, Georgetown University and (b) CT suite, Philips Research.
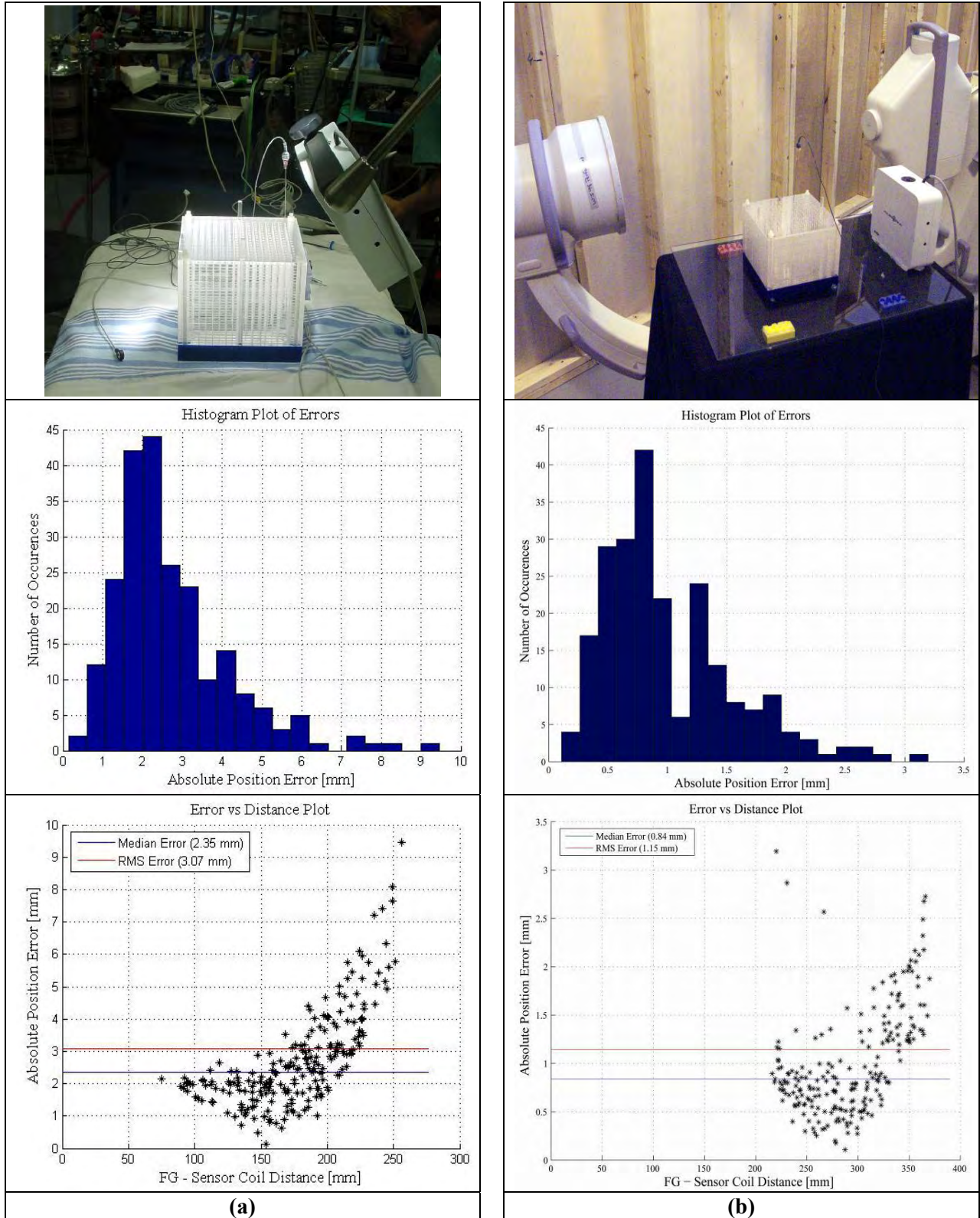
Figure 4: Summary of results within (a) Animal OR suite, CSTAR and (b) C-arm test suite, GE Research.

4. Discussion and Conclusions

It is hoped that this work brings to the fore the need for a common hardware phantom and analysis methodology for evaluating the accuracy of electromagnetic tracking in the clinical environment. We have presented both a hardware phantom and analysis methods whose design is based on our experience in the use of electromagnetic tracking in the clinical environment. Both the hardware and an implementation of the analysis methods were used to assess the accuracy of electromagnetic tracking at four institutes in various laboratory and clinical settings.

A curious difference was the variability in baseline results at different institutes. From the tests at Georgetown University it is clear that while every effort was made to remove any visible metallic objects from a one meter radial distance from the test area, comparison of baseline results indicate that distortion from unavoidable sources such as computers and monitors, overhead lights and wiring conduits within the building, while further than two meters away, still induce a significant amount of distortion within the electromagnetic measurement field. Surprisingly, the DynaCT suite, a clinical environment, yielded the best accuracy in the tests at Georgetown. This may be because the patient table is made of carbon fiber and is suspended further away from the electronic and metal components used to position the couch. In conventional CT and OR suites, the couch positioning components typically rest directly below the patient. Therefore, raising the measurement volume and field generator further up is one of the few changes that can be made. However, the diameter of the CT gantry will place an upper limit on how high the patient location can be raised while still being clinically feasible.

As was expected from the onset of this study, the accuracy results vary widely between clinical environments. However, it is to be noted that the accuracy of electromagnetic tracking systems appears to be within acceptable bounds for a large variety of clinical applications. With a better understanding of the distortion factors in each environment, applications that rely on electromagnetic trackers for navigation can be optimized to work efficiently provided necessary steps are taken. In our experience these necessary step include positioning the field generator and/or measurement volume further away from the distortion source and in rare cases might involve making a change in the clinical protocol such that it is commensurate with the change in location of measurement volume and field generator placement. Given the continued improvement in the measurement accuracy and resilience of electromagnetic tracking systems we expect their applicability within the clinical environment to increase over the coming years. As electromagnetic tracking systems improve in accuracy and better address some of these distortion issues, it will become less important to modify and shape the clinical protocol to facilitate incorporation of electromagnetic trackers.

A significant shortcoming of this iteration of the hardware phantom is the use of several layers to construct the cube. The non-uniform thickness of each layer necessitated the measurement of each hole depth using a depth gauge. In addition, despite the four corner rods that secure the layers, they have a tendency to slide under impact. Some of these mechanical limitations of the initial iteration of the hardware phantom will be addressed and assuaged with the next iteration that will use as few as four layers, wherein each layer is extruded from a thicker Acrylic sheet. Thus uniform layer thickness can be assured and layer slippage can be minimized. Another issue to be addressed as part of this ongoing study is the selection of registration nodes. The ideal solution would be to compute statistics for all possible choices of registration nodes and then select the eight nodes that together provide the lowest error statistic. This however, is far from feasible due to the inordinate length of time taken to cycle through all possible combinations. A possible solution is to select a subset of the entire node set and recursively search through this smaller subset for best registration node choices.

Another shortcoming of the protocol and phantom is the lack of testing of the rotational accuracy of the five degrees of freedom sensors. Here the analysis has been limited to the positional accuracy but the 5D sensors do provide two degrees of freedom for the sensor's orientation. In future work, a set of directional statistics[17] will be included that reflect the rotational accuracy of the sensor. Further, we will investigate phantom designs in which the set of needle holes are not parallel to one another.

Based on these results and results form previous studies at our institute (Georgetown University Medical Center), we feel confident that electromagnetic tracking systems provide sufficient accuracy to be considered for clinical trials. To this end, we are currently pursuing an Institutional Review Board (IRB) human trial to use electromagnetic tracking for image-guided biopsy procedures within the abdominal cavity.

Acknowledgements

References

[1] J. B. Anon "Computer-Aided Endoscopic Sinus Surgery" Laryngoscope, July 1998;108 (7): 949-961.

[2] F Watzinger, W Birkfellner, F Wanschitz, et al., "Positioning of Dental Implants Using Computer-Aided Navigation and an Optical Tracking System: Case Report and Presentation of a New Method" J Craniomaxillofacial Surg, April 1999; 27 (2): 77-81.

[3] P. G. Seiler, H Blattmann, S Kirsch, et al., "A Novel Tracking Technique for the Continuous Precise Measurement of Tumor Positions in Conformal Radiotherapy" *Physics in Medicine and Biology,* 2000*, 45*: N103-N110.

[4] M Fried, J Kleefield, H Gopal, et al., "Image-Guided Endoscopic Surgery: Results of Accuracy and Performance in a Multicenter Clinical Study Using an Electromagnetic Tracking System", Laryngoscope, May 1997; 107(5): 594-601.

[5] M. A. Nixon, B.C.McCallum, W. R. Fright and N. B. Price, "The Effects of Metals and Interfering Field on Electromagnetic Trackers" *Presence,* 1998*, 7 (2)*: 204-218.

[6] V. V. Kindratenko "A Survey of Electromagnetic Position Tracker Calibration Techniques" VR: Research, Development and Applications, 2000, 5 (3): 169-182.

[7] W Birkfellner, F Watzinger, F Wanschitz, et al., "Systematic Distortions in Magnetic Position Digitizers" Med Phys, 1998, 25 (11): 2242-2248.

[8] J Hummel, M. R. Bax, M. L. Figl, et al., "Design and Application of an Assessment Protocol for Electromagnetic Tracking Systems" Med Phys, 2005, 32 (7): 2371-2380.

[9] J Hummel, C. M. Jun, M, Figl, et al, "Standardized Evaluation Method for Electromagnetic Tracking Systems" SPIE: Medical Imaging, 2005, 5744: 236-240.

[10] Z Yaniv and K Cleary, "Fluoroscopy Based Accuracy Assessment of Electromagnetic Tracking" SPIE: Medical Imaging, 2006; 6141: 168-174.

[11] E Wilson, R Slack, F Banovac, et al., "Electromagnetic Tracker Accuracy in the CyberKnife Suite" SPIE: Medical Imaging, 2006; 6141:61411R.

[12] C.A. Nafis, V. Jensen, L. Beauregard, P.T. Anderson, "Method for Estimating Dynamic EM Tracking Accuracy of Surgical Navigation Tools", SPIE: Medical Imaging, 2006; 6141: 152-167.

[13] B. J. Wood, H Zhang, A Durrani, et al., "Navigation with Electromagnetic Tracking for Interventional Radiology Procedures: A Feasibility Study" J Vasc. Radiol., 2005; 16: 493-505.

[14] F. Banovac, J Tang, S Xu, et al., "Precision Targeting of Liver Lesions Using a Novel Electromagnetic Navigation Device in Physiologic Phantom and Swine" Med. Phys., 2005; 32 (8): 2698-2706.

[15] F Banovac, E Wilson, H Zhang and K Cleary, "Needle Biopsy of Anatomically Unfavorable Liver Lesions with an Electromagnetic Navigation Assist Device in a Computed Tomography Environment" J Vasc. Interv. Radiol., Oct 2006; 17 (10): 1671-1675.

[16] http://www.ndigital.com/aurora-techspecs.php

[17] D. D. Frantz, A. D. Wiles, S. E. Leis and S. R. Kirsch "Accuracy Assessment Protocols for Electromagnetic Tracking Systems" Phys. Med. Biol., 2003; (48): 2241-2251.

[18] B.K.P. Horn, "Closed-form solution of absolute orientation using unit quaternions", Journal of the Optical Society of America, 4(4), 629-642, 1987.

[19] K. Gary et al., "IGSTK: An Open Source Software Toolkit for Image-Guided Surgery", IEEE Computer, 39(4), 46-53, 2006.

8.7 Wong 2006: Creation of 4D Imaging Data …

Reprint begins on the next page and is ten pages.

Creation of 4D Imaging Data using Open Source Image Registration Software

Kenneth H. Wong[*a], Luis Ibanez[b], Teo Popa[a], and Kevin Cleary[a]

[a]ISIS Center, Department of Radiology, Georgetown University, Washington DC

[b]Kitware, Clifton Park, NY

ABSTRACT

4D (3 spatial dimensions plus time) images using CT or MRI will play a key role in radiation medicine as techniques for respiratory motion compensation become more widely available. Advance knowledge of the motion of a tumor and its surrounding anatomy will allow the creation of highly conformal dose distributions in organs such as the lung, liver, and pancreas. However, many of the current investigations into 4D imaging rely on synchronizing the image acquisition with an external respiratory signal such as skin motion, tidal flow, or lung volume, which typically requires specialized hardware and modifications to the scanner. We propose a novel method for 4D image acquisition that does not require any specific gating equipment and is based solely on open source image registration algorithms. Specifically, we use the Insight Toolkit (ITK) to compute the normalized mutual information (NMI) between images taken at different times and use that value as an index of respiratory phase. This method has the advantages of (1) being able to be implemented without any hardware modification to the scanner, and (2) basing the respiratory phase on changes in internal anatomy rather than external signal. We have demonstrated the capabilities of this method with CT fluoroscopy data acquired from a swine model and cross-sectional imaging data from human volunteers.

Keywords: CT, MRI, dynamic imaging, radiation therapy, open-source software

1. INTRODUCTION

1.1. Clinical challenges in conformal radiation therapy

In radiation therapy, the overarching goal is to deliver a lethal dose to the cancerous tissue while minimizing collateral damage to the surrounding normal tissues. Imaging plays a key role in this effort because the dose distribution can only be as accurate as our ability to define what is the target and what is normal tissue. Most treatment planning therefore uses cross-sectional imaging (primarily CT, although MRI and PET/SPECT can also be used) to accurately determine the location and shape of the tumor. These images contain features (e.g. bony anatomy) or artificial landmarks that can be identified immediately before the patient receives a radiation treatment so that the technologist can ensure that the patient is in the correct position and that the radiation dose will be delivered according to the treatment plan. An inherent assumption in this process is that the patient anatomy has not appreciably changed between the time of treatment planning and the time that the treatment takes place. Although this assumption is generally true in bony regions such as the skull and spine, it is generally false in the thorax and abdomen, where respiration causes much of the anatomy to be in constant motion and deformation. For example, a review of literature on respiratory motion of the liver found that the displacement during quiet respiration is in the range of 10-26 mm, and that deep breathing could drive this range as high as 75 mm[1]. Liver deformation measured using breath-hold MRI and a statistical shape model showed overall deformations in the range of 3-4 mm and local deformations as high as 23 mm[2]. Furthermore, breathing patterns can be irregular in amplitude and frequency[3,4], and can vary greatly from patient to patient. Thus, any treatments in the abdomen and thorax based on a single static cross-sectional image, such as a breath-hold CT scan, are potentially subject to significant targeting errors.

1.2. Methods for imaging respiratory motion

Clearly, if we wish to achieve the highest conformality for radiation dose delivery in moving organs, we must have a means to measure and model these changes. Accordingly, there has been significant interest in developing

---

[*] wong@isis.georgetown.edu; phone 202-784-1521; fax 202-784-3479; www.isis.georgetown.edu

methodologies for dynamic or 4D imaging techniques that can measure the changes that occur during respiration. Typically, these 4D images are a series of standard 3D CT volumes acquired from different points in the respiratory cycle (e.g. end-inspiration, end-exhalation, and various times in between). From a technical standpoint, the simplest approach would be to instruct patients to hold their breath several times and acquire several CT scans over the same volume. However, it is very challenging for people to hold their breath at anything other than full exhalation or full inspiration, so this approach is not very practical. Instead, the usual approach is to place an external marker on the skin (such as an optically tracked beacon) and measure the motion of this marker. The external respiratory signal can then be used to either trigger the acquisition of the CT image (Figure 1) or to serve as a "timestamp" for when the image was acquired (Figure 2).
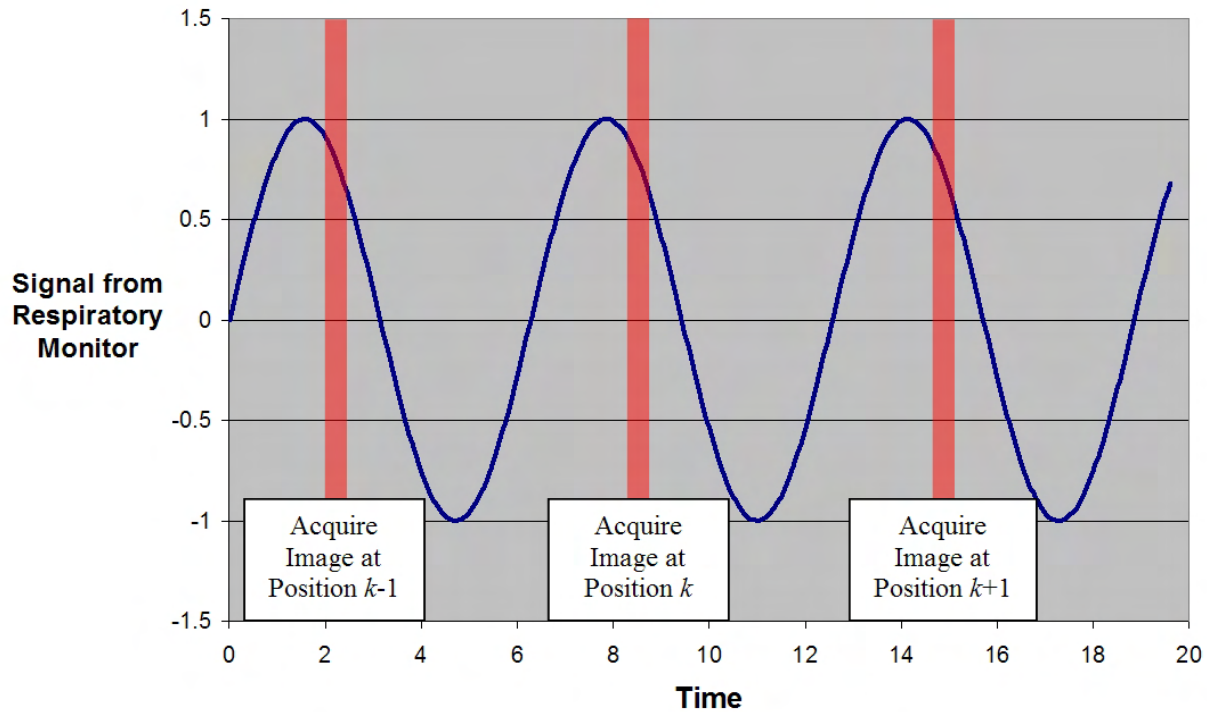

Figure 1: Illustration of 4D image acquisition using a prospective gating approach. The sinusoidal trace is the input from a respiratory monitoring system such as an external skin marker. Immediately after full exhalation, the monitoring system triggers image acquisition at one axial position. By triggering image acquisition (vertical bars) to the signal from the external skin marker, we can acquire a full volume representing one segment of the respiratory cycle. To obtain the next segment of the respiratory cycle, we would simply adjust the triggering parameters and repeat the acquisition.

Previous studies of 4D imaging techniques have focused on the latter implementation, where several image sets of the target volume at one couch position are acquired under free breathing. The couch is then shifted and the image acquisition is repeated until the desired patient volume is covered. Once the volume has been completely covered, the images are sorted into bins corresponding to the different phases of the respiratory cycle, and a full 4D volume can be assembled from these images. Ford et al. used an external respiration monitor to mark images from a single slice CT scanner with a corresponding respiratory phase, and then sorted the images into different volumes to form a complete 4D image set[5]. Vedam et al. and Pan et al. used a similar experimental setup with multislice CT scanners[6-8]. The advantage of the multislice scanner approach is obviously that a larger volume can be acquired during a single gantry rotation, but all of the methods share the same basic approach.
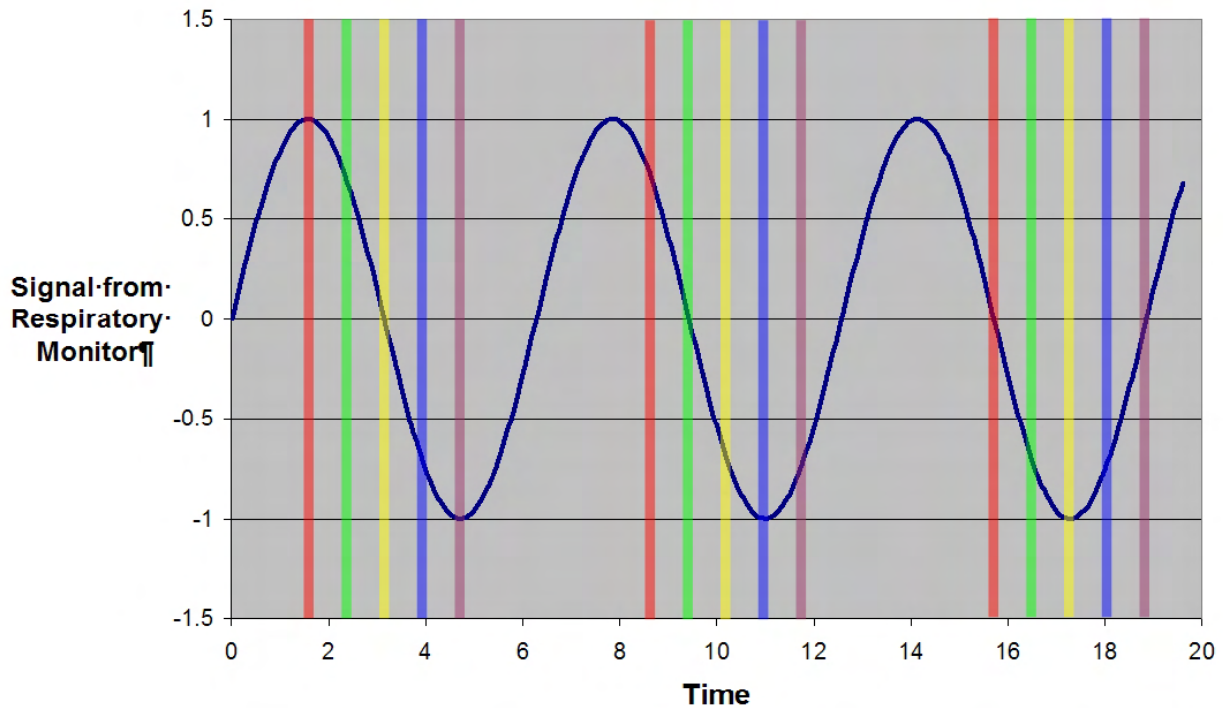
Figure 2: Schematic for retrospective respiratory gating with image sorting. Images are acquired during free breathing and each image is timestamped with a respiratory signal. Each vertical bar represents a single acquisition, so on a multislice scanner each bar could represent several axial slices. Once the images have been collected, they are sorted based on the timestamp. A complete set of images from each axial position in the patient then covers the entire respiratory phase, and the total set of images is the 4D volume.

1.3. Open source image registration approach for 4D imaging

It is clear that there is both a strong clinical need for 4D imaging and a significant effort underway to develop methods for 4D imaging. Our goal in this project is to develop a purely open source software methodology for 4D image acquisition. An open source approach provides the ability to easily disseminate the software methods and allow many other investigators to use and test the methods with their own image data. A pure software approach also eliminates much of the variability associated with acquisition and triggering hardware in different scanner manufacturers and scanner models. Removing that barrier will allow these methods to be implemented on a widespread basis.

2. METHODS

2.1. 4D Imaging data space

The general 4D image acquisition process can be thought of as filling the data space illustrated in Figure 3, where the horizontal axis is the table or axial position and the vertical axis is the phase of the respiratory cycle in radians. Each symbol on the figure represents a single axial 2D slice. In this figure, the symbols illustrate an idealized 4D acquisition that is evenly spaced in both the spatial and the phase domains. A conventional CT scan taken during breath hold would therefore be represented as a horizontal line across the data space. The prospective gating approach shown in Figure 1 would fill up a single line of this data space, and then the triggering parameter would need to be adjusted in order to acquire the next line. Likewise, a vertical line through the space would represent a series of images taken at a single table position at different points in the respiratory cycle, such as CT fluoroscopy or cine MRI. A retrospective gating approach with a spiral CT scanner (as shown in Figure 2) would look like a diagonal line drawn through the

space, because during the acquisition of the data the patient is breathing and the scanner is simultaneously moving the patient through the gantry.
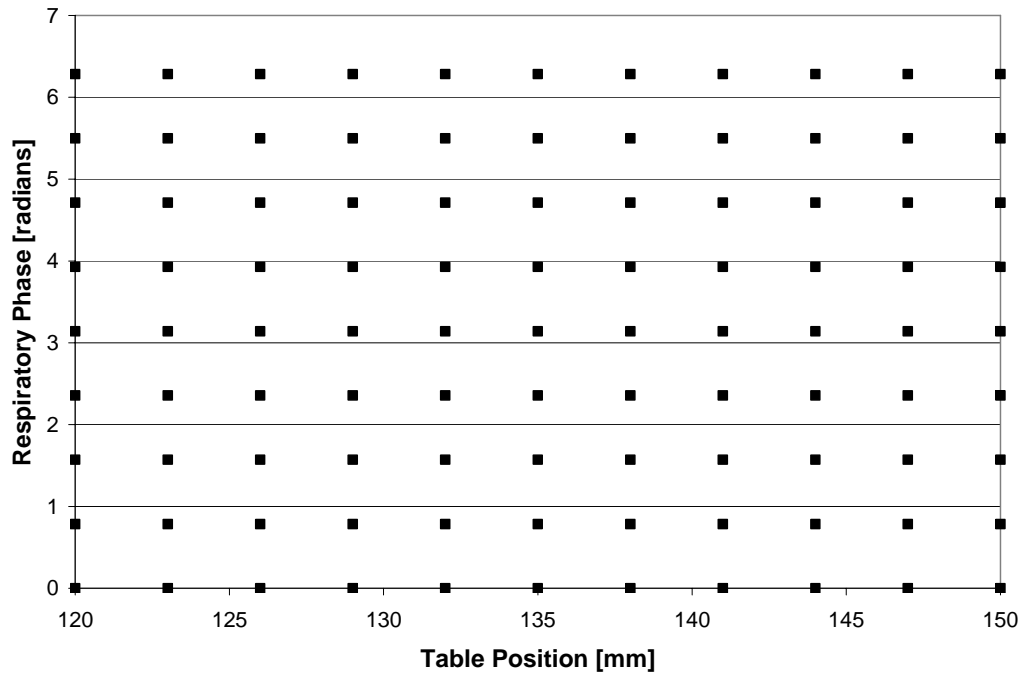

Figure 3: Data space for 4D image acquisition. The horizontal axis represents a table position (assuming transverse image slices, as used in CT) and the vertical axis represents the phase in the respiratory cycle. In this diagram, each symbol represents one image in a 4D data set, so the ensemble of symbols covers an entire volume and all times within the respiratory cycle.

The power of this conceptual diagram is that any image from a patient falls somewhere in this space, so no matter how we acquire images from the patient, we can use this diagram to classify and sort them. The table position or the axial location of the image is known by the scanner, so the positioning of the symbol along the horizontal axis is straightforward; the challenge in 4D acquisition is to assign each image a phase within the respiratory cycle.

2.2. Image comparison method for 4D image sorting

Previous studies in 4D imaging have obtained the respiratory phase information by monitoring an external signal and using this information to timestamp the images and sort them into bins. Our approach is to derive this timestamp from comparing the images to each other or to a reference image. Specifically, we use ITK (the Insight Toolkit, an open source toolkit for image segmentation and registration) to calculate the normalized mutual information (NMI) between the first image of the series and all successive images. Since respiratory motion can be thought of as a continuous periodic deformation of the anatomy, this NMI should be periodic and therefore its value is a measure of the phase of each image.

2.3. Testing with CT fluoroscopy data - measurement of respiratory period

We have tested this method on CT fluoroscopy data from swine, which were being imaged under an approved animal protocol during testing of a robotic lung biopsy system (see Paper 6141-57 in these Proceedings). The animal was anesthetized, intubated, and placed on a mechanical ventilator. For these experiments, the animal ventilator was set to 11 breaths per minute, which translates to a period of 5.4 seconds. CT fluoroscopy images were acquired at a rate of 4

Hz with a 512 x 512 image matrix and a slice thickness of 5 mm using a tube voltage of 140 kvP and a tube current of 100 mA. At each table position, CT fluoroscopy data were acquired over at least 2 full respiratory cycles to generate a cine sequence. DICOM CT fluoroscopy images were stored and then analyzed offline. For each table position, we used the first CT fluoroscopy image in the sequence as the reference image. We then used ITK to calculate the NMI between the reference image and successive images in the sequence.

2.4. Testing with 4D_CT data from swine model - identification of respiratory phase

Previous work by Xu *et al.* in our group has produced 4D CT data from a swine model based on a custom image registration approach[9,10]. Image data were acquired using a similar protocol to the one described in the previous section, and the lung fields were segmented from the image for use in the registration. In order to convert the CT fluoroscopy data into 4D CT data, we subdivided each single CT slice into small overlapping subregions and registered the subregions against a previously obtained reference CT volume by minimizing the Zero Mean Sum of Squared Differences (ZSSD). A quadratic registration was used to model the deformation between the two regions, and any inter-modality differences between the standard CT and the CT fluoroscopy images are corrected by normalizing the mean intensities of the images. We derived an overall displacement based on the effect of many local registrations of small regions. The average of all the pixel displacements then yielded a 3D motion vector for each CT fluoroscopy slice, which reflects the overall phase of the slice within the respiratory cycle. Using principal component analysis, we then determined the major axis of the motion trajectory. The component of each 3D motion vector along that major axis is essentially a metric of the respiratory phase, which allows us to assemble the 2D fluoroscopy slices into a 4D volume.

These data represent a known source of 4D image data covering the entire data space from Figure 3, and thus provides another test for our methodology. We extracted slices from the known 4D volume and then used NMI comparisons to determine if we could correctly identify the phase of the extracted images relative to a 3D volume obtained at end-exhalation (this volume was obtained by briefly stopping the ventilator).

3. RESULTS

3.1. CT fluoroscopy data - determination of respiratory period

Figure 4 shows a plot of the results from one CT fluoroscopy series taken at a single fixed table position. The horizontal axis is the acquisition time and the vertical axis is the NMI between the first image of the series and the successive images. We observe that the NMI fluctuates periodically as the animal breathes, and that the period for this fluctuation is approximately 5.5 seconds. This signal also looks very much like what might be obtained from a respiratory belt or optical tracker monitoring the skin, albeit somewhat more noisy.

We also used Matlab to calculate the power spectrum of the signal from Figure 4, based on the signal's FFT.

Figure 5 clearly shows a single dominant frequency, which is at 0.1875 Hz, corresponding to a breathing period of 5.3 seconds. Taken together, these results show that this method can correctly identify the respiratory period of a breathing animal and can be used to monitor respiration and sort a series of axial images into bins based on their location in the respiratory cycle.

3.2. CT swine data - identification of respiratory phase

For this test, a 3D volume taken at one point of the respiratory cycle (in this case, end-exhalation) represents a reference for comparison against the 4D volume. If we compare the slices in the 4D volume to the slices in the 3D volume using NMI (while keeping the table position constant), we should find that the NMI value oscillates according to the respiratory phase. Furthermore, the minimum value of NMI should occur when the respiratory phase is exactly opposite that of the reference volume. Thus, if our reference 3D scan is at end-exhalation, we should observe the minimum NMI value at end-inspiration.

Results from these calculations are shown in Table 1 and Figure 6. Each column in the table represents one table/axial position in the 4D volume, and each entry within the column is the NMI between a slice from the 4D volume and a slice from the 3D reference volume. The shaded boxes in each column represent the location of end-inspiration as

determined from the original 4D volume; these slices correspond very well to the location of the minimum value of the NMI. We conclude that the NMI can be used to identify respiratory phase and thereby sort images in a 4D volume.
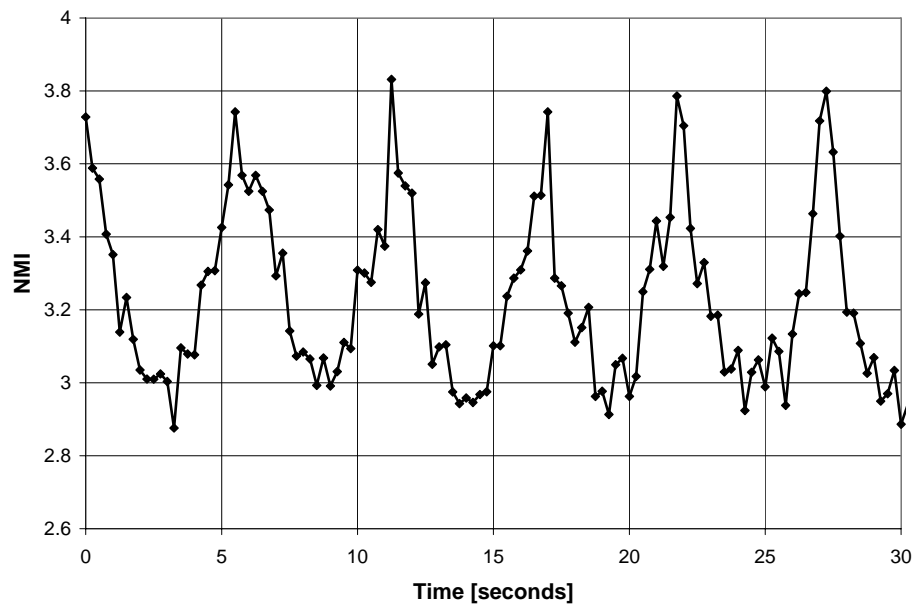

Figure 4: Plot of Normalized Mutual Information (NMI) versus time for swine data. The respiratory period is approximately 5.5 seconds, which corresponds well with the ventilator setting from the animal experiment (approximately 11 breaths/minute).
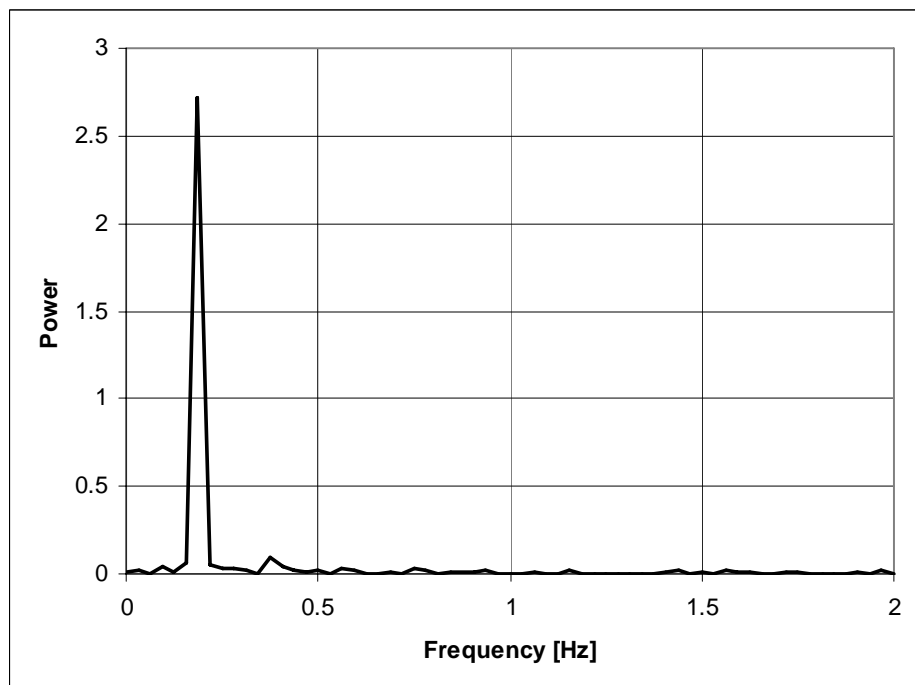
Figure 5: Power spectrum for the signal shown in Figure 4. There is a clearly dominant frequency at 0.1875 Hz, which corresponds to a breathing period of 5.33 seconds. This number is also in good agreement with the ventilator setting from the animal experiments.

Image Number	Table/Axial position							
	0	1	2	3	4	5	6	7
1	0.509	0.700	0.787	0.729	0.494	0.685	0.687	1.02
2	0.501	0.680	0.738	0.645	0.479	0.706	0.668	0.843
3	0.497	0.610	0.591	0.628	0.482	0.719	0.530	0.789
4	0.519	0.569	0.595	0.605	0.498	0.723	0.505	0.751
5	0.520	0.546	0.615	0.623	0.504	0.756	0.473	0.776
6	0.523	0.547	0.614	0.628	0.511	0.766	0.480	0.780
7	0.555	0.564	0.622	0.639	0.520	0.805	0.493	0.808
8	0.559	0.551	0.639	0.639	0.541	0.806	0.498	0.816
9	0.553	0.555	0.647	0.651	0.545	0.849	0.514	0.832
10	0.574	0.596	0.644	0.672	0.544	0.897	0.525	0.858
11	0.608	0.602	0.643	0.681	0.577	0.910	0.531	0.910
12	0.626	0.592	0.721	0.693	0.588	0.947	0.544	0.926
13	0.593	0.598	0.677	0.728	0.586	0.978	0.553	0.952

Table 1: Values of NMI calculated between a 4D volume and a reference 3D volume obtained at end-exhalation. The values are calculated for a particular table position as listed by columns. Each entry in the column is the NMI between a 2D slice from the 4D volume and a 2D slice from the 3D volume. Shaded boxes indicate the location of end-inspiration, as determined from analysis of the 4D volume, and correspond with the minimum value of the NMI.
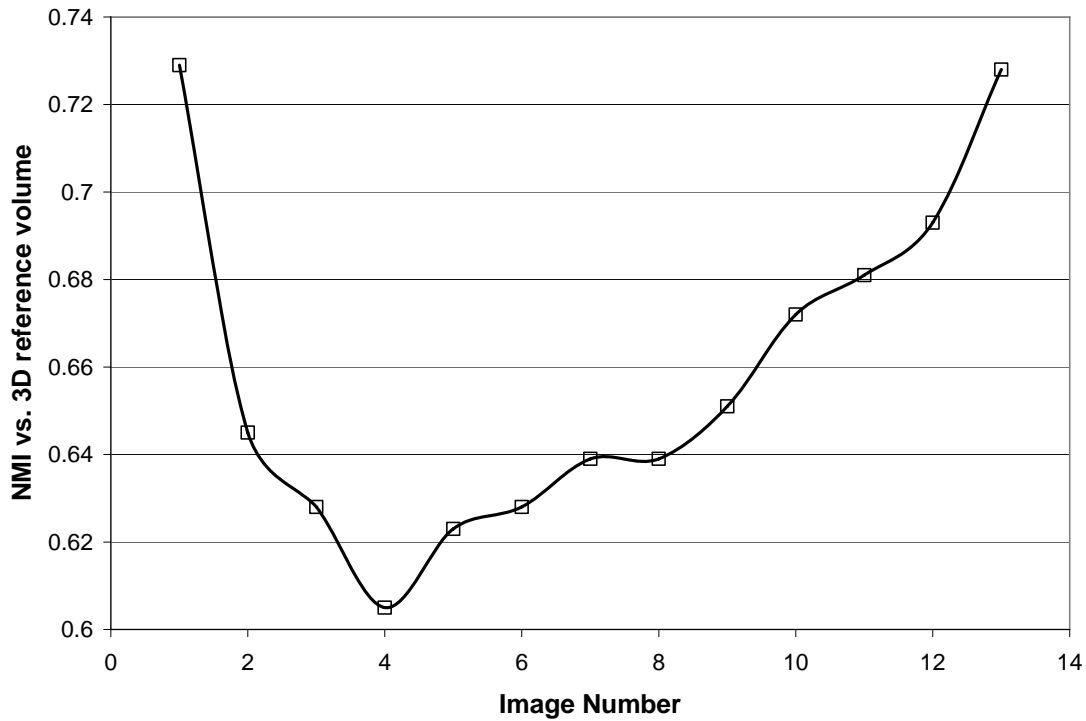
Figure 6: Plot of the NMI values from Table Position 3. The NMI oscillates and reaches its minimum at image number 4, which corresponds to end-inspiration in the 4D volume.

4. DISCUSSION

To our knowledge, this is the first time that a purely software/image processing approach to 4D imaging has been presented. 4D imaging without an external respiratory signal is novel, and could have distinct advantages over the current methods. First, it is easier to implement because it does not require any external respiratory gating hardware to interface with the CT or MRI scanner. Second, it measures respiratory phase based on internal motion rather than external motion. Since the end goal of respiratory motion compensation is to be able to localize internal tumors, it is advantageous to directly track internal motion rather than track external motion and infer an internal location. Finally, NMI or some other relatively simple image metric is easy to calculate, so processing 4D data does not require long computation times that might be required using a multi-step deformable registration approach as used by Xu[10].

This method shares some similarity with recent work in 4D CT by Pan *et al.*[8]. where the summed Hounsfield Units in a region of interest were used as an index of respiratory phase. However, computation of the NMI is more objective and faster to implement because it does not require the user to choose a region of interest. Furthermore, the computation of summed Hounsfield Units in a region of interest is essentially evaluating a subset of the image histogram, whereas our method uses the entire image histogram.

In our current implementation, we use the NMI values and the known table positions to sort 2D images into a 4D volume. An alternate approach would be to use the acquired data as samples of the data space from Figure 3, which can be thought of as a continuum that covers the range of possible table positions and the full extent of the respiratory cycle. Given a sufficient number of samples covering the data space, it should be possible to interpolate a 2D slice at any chosen table position and respiratory phase. Obviously, this interpolation will be complex because we are interpolating between images rather than numbers, but such an approach might allow for newer and more efficient ways to acquire the 4D volume. We may draw a parallel here with MRI research into novel techniques for k-space filling, which take advantage of symmetry and Fourier transform properties in order to optimize the image acquisition[11].

Based on the above concept, one potential area for future investigation is to determine the sampling necessary to create a full 4D imaging volume, i.e., what is a sufficient sampling density in the space to be able to recover any arbitrary 2D slice? The answer to this is likely to be dependent on the information that we seek from our images. For example, if high-contrast fiducial markers are implanted near or inside the tumor, it may be possible to use a much coarser sampling of the data space because the features of interest in the image are easier to recognize. Likewise, if we need to visualize fine soft tissue detail from the images, a denser sampling may be required. Addressing this question is especially important in 4D CT, where each acquired image carries a nonzero risk from radiation exposure and there will be an increasing desire to acquire 4D CT data for radiation therapy treatment planning.

5. CONCLUSIONS

We have developed and tested a novel method for 4D imaging that is implemented using open source software. This method should be applicable to many types of medical images, can be used without any specialized gating hardware or modifications to the scanner, and is also vendor-independent. In future work, we will continue to test the algorithm with different data sources in order to verify its robustness and applicability to the arena of respiratory motion compensation in radiation medicine. We will also investigate whether or not the 4D imaging capability can be used to improve respiratory motion compensation systems used in CyberKnife radiosurgery treatments.

ACKNOWLEDGMENTS

REFERENCES

1. M.A. Clifford, F. Banovac, E. Levy, and K. Cleary, "Assessment of hepatic motion secondary to respiration for computer assisted interventions", *Comput Aided Surg* 7 (5), p291, 2002.

2. J. M. Blackall, "Respiratory Motion in Image-Guided Interventions of the Liver", Doctoral Thesis, King's College, 2002.

3. G .C. Donaldson, "The chaotic behaviour of resting human respiration", *Respir Physiol* 88 (3), p313, 1992.

4. E. N. Bruce, "Temporal variations in the pattern of breathing", *J Appl Physiol* 80 (4), p1079, 1996.

5. E. C. Ford, G. S. Mageras, E. Yorke, C. C. Ling, "Respiration-correlated spiral CT: a method of measuring respiratory-induced anatomic motion for radiation treatment planning." *Med Phys* 30(1): 88-97, 2003.

6. Vedam SS, Keall PJ, Kini VR, Mostafavi H, Shukla HP, Mohan R, "Acquiring a four-dimensional computed tomography dataset using an external respiratory signal," *Phys Med Biol.* Jan 7;48(1):45-62, 2003.

7. P. Keall, "4-dimensional computed tomography imaging and treatment planning." *Semin Radiat Oncol* 14(1): 81-90, 2004.

8. T. Pan, T. Y. Lee, E. Rietzel, G. T. Chen, "4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT." *Med Phys* 31(2): 333-40, 2004.

9. Xu, S., Fichtinger, G., Taylor, R. H. and Cleary, K., "3D motion tracking of pulmonary lesions using CT fluoroscopy images for robotically assisted lung biopsy". SPIE Medical Imaging Conference, San Diego, CA, 2005.

10. Xu, S., Taylor, R. H., Fichtinger, G. and Cleary, K.,"Lung Deformation Estimation and Four-dimensional CT Lung Reconstruction". Medical Image Computing and Computer-Assisted Intervention, Palm Springs, CA, USA, Springer, 2005.

11. McRobbie DW, Moore EA, Graves MJ, Prince MR, *MRI: From Picture to Proton*, Cambridge University Press, Cambridge, UK, 2003, p.317-323.

8.8 Yaniv 2007: The Influence of CT Based …

Reprint begins on the next page and is seven pages.

The influence of CT based attenuation correction on PET/CT registration: an evaluation study

Ziv Yaniv, Kenneth H. Wong, Filip Banovac, Elliot Levy, Kevin Cleary

Imaging Science and Information Systems (ISIS) Center, Dept. of Radiology,
Georgetown University Medical Center, Washington, DC, USA.

ABSTRACT

We are currently developing a PET/CT based navigation system for guidance of biopsies and radiofrequency ablation (RFA) of early stage hepatic tumors. For these procedures, combined PET/CT data can potentially improve current interventions. The diagnostic efficacy of biopsies can potentially be improved by accurately targeting the region within the tumor that exhibits the highest metabolic activity. For RFA procedures the system can potentially enable treatment of early stage tumors, targeting tumors before structural abnormalities are clearly visible on CT. In both cases target definition is based on the metabolic data (PET), and navigation is based on the spatial data (CT), making the system highly dependent upon accurate spatial alignment between these data sets. In our institute all clinical data sets include three image volumes: one CT, and two PET volumes, with and without CT-based attenuation correction. This paper studies the effect of the CT-based attenuation correction on the registration process. From comparing the pairs of registrations from five data sets we observe that the point motion magnitude difference between registrations is on the same scale as the point motion magnitude in each one of the registrations, and that visual inspection cannot identify this discrepancy. We conclude that using non-rigid registration to align the PET and CT data sets is too variable, and most likely does not provide sufficient accuracy for interventional procedures.

Keywords: Image-Guided Therapy, PET/CT, non-rigid registration, PET attenuation correction

1. INTRODUCTION

We are currently developing a PET/CT based navigation system for guidance of biopsies and Radio Frequency Ablation (RFA) of hepatic tumors.[1] For RFA procedures combined PET/CT data can potentially enable treatment of tumors before structural abnormalities appear on CT, or when these abnormalities are not clearly visible. For tumor biopsies this system can potentially improve diagnostic results by accurately targeting regions within the tumor that exhibit the highest metabolic activity. In both cases we use PET data to define a spatial target which is not clearly visible on the CT. The underlying assumption of our approach is that the PET and CT are spatially aligned.

To date, combined Positron Emission Tomography (PET) and Computed Tomography (CT) machines have been primarily used for diagnostic purposes in the oncological setting. They have led to improved diagnostic accuracy,[2] and higher patient throughput. The former is due to the fusion of spatial and functional information, while the later is due to the higher speed of CT-based attenuation correction as compared to PET transmission based correction.

These combined machines facilitate image fusion by providing intrinsically registered data sets, under the assumption that there is a known *rigid* relationship between the PET and CT data. In the regions of the abdomen and thorax this is an approximation as the data sets are related via a non-rigid transformation. This inconsistency is a result of the acquisition protocols. CT images are acquired at breath-hold while PET images are acquired with normal tidal breathing, due to the long acquisition times (several minutes per slice).

The visual quality and quantitative accuracy of PET data is primarily influenced by photon attenuation, necessitating the use of attenuation correction schemes.[3] One possible correction scheme is based on the use of CT data for X-ray based attenuation correction, and is the method used in combined PET/CT scanners.[4]

E-mail: zivy@isis.georgetown.edu
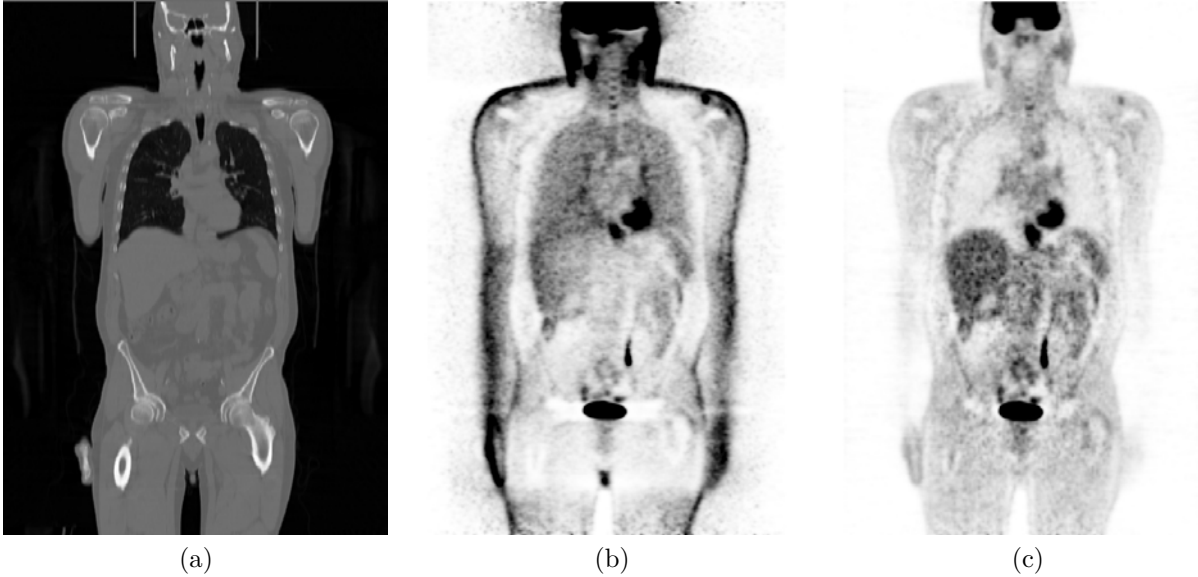
(a)   (b)   (c)

Figure 1. Corresponding coronal slices from a combined PET/CT machine: (a) CT, (b) PET without attenuation correction (c) PET after CT-based attenuation correction. Dark regions in the PET data correspond to high metabolic activity. Note the typical enhanced activity in the region of the lungs prior to attenuation correction.

The primary assumption underlying this scheme is that the spatial relationship between the PET and CT coordinate systems is known (not necessarily rigid). Currently, most PET/CT machines assume that this spatial relationship is both known and rigid. This assumption is invalid in the regions of the abdomen and thorax, resulting in attenuation correction errors. To mitigate the effect of these errors on diagnosis most PET/CT machines output three corresponding data sets: CT, and PET with and without attenuation correction, such as those shown in Figure 1. Evaluating the results is thus deferred to the physicians, relying on their ability to identify imaging artifacts due to respiratory motion.

Thus, combined PET/CT machines produce PET data that is spatially inconsistent with the known CT coordinate system, which in turn leads to erroneous attenuation correction.

A comprehensive solution to these problems is to acquire 4D (3D over time) CT and PET data sets. This approach is based on retrospective or prospective gating. In[5] CT is gated retrospectively and PET is gated prospectively. A respiratory signal is obtained from an optically tracked marker placed on the patient's chest. To acquire 4D CT the respiratory cycle is divided into bins and the CT slices are time-stamped with their respective phases in the respiratory cycle using a signal triggered by the imaging apparatus indicating slice acquisition. The resulting set of binned slices represent the CT volume over time. To acquire a 4D PET data set a prospective gating approach is used with data acquisition triggered by the respiratory signal. In this case the activation events are binned such that each bin represents a single 3D PET volume. A similar approach was used in.[6] In this work both CT and PET data were gated retrospectively, and the respiratory signal was based on the temperature of the patient's breathing airflow.

A possible problem with gated PET reconstruction is the small number of activations per bin. To obtain enough data per bin for a high quality reconstruction longer acquisition times and increased amounts of radioactive tracer material are required, both of which are preferably avoided. A recently proposed solution to this problem is to non-rigidly register all low quality PET volumes to a single phase,[7] resulting in a single high quality volume.

While these are comprehensive solutions they require interaction with the imaging apparatus, either to trigger acquisition or to receive a signal that data has been acquired. These hardware interfaces are not always available which has lead researchers to make the following implicit assumption: CT and PET data sets are correct

representations of the underlying anatomical structures for two different points in the respiratory cycle, and can thus be aligned using non-rigid registration.[8, 9]

Currently we do not have access to the internal workings of the imaging apparatus . As our navigation system assumes that PET and CT data are spatially aligned we turn to a registration-based approach, which in many cases is a sufficiently accurate approximation.

In our institute all clinical data sets include three image volumes, CT, and two PET data sets, before and after CT-based attenuation correction. As both PET data sets are spatially equivalent the transformation obtained by non-rigid PET/CT registration should be approximately the same for both PET data sets. The purpose of the work reported here is to assess the effect of the CT-based attenuation correction on non-rigid registration, and to evaluate the reliability of visual inspection as a qualitative method for evaluating non-rigid PET/CT registration when ground truth is not available.

2. METHODS

Non-rigid PET/CT registration has been previously addressed using a variety of approaches including the use of free-form deformations[9] and hierarchical piecewise rigid registration.[8, 10] It has been demonstrated that deformable registration of CT-based attenuation corrected PET and CT can potentially improve the alignment between the two data sets.[10]

We chose to use the Free Form Deformation (FFD) based registration method described in[9] and implemented in the Insight Toolkit (ITK).[11] Registration is cast as an optimization task with the objective function defined by image similarity. In our case images are related via a stochastic process and mutual information serves as the similarity measure. To deform the image volume it is embedded inside a grid of node points that control a set of approximating B-splines. The volume is then deformed by changing the position of the node points. Our choice of the FFD based registration method is due to its successful record in registering various anatomical structures including the liver[12] and its previous use for PET/CT registration of the chest region.[9]

All data was acquired using a Siemens Biograph 6 PET/CT machine. The original data are whole-body scans with the following image resolutions: $512 \times 512 \times 244$, for CT, with a voxel size of $1.3 \times 1.3 \times 4mm^3$, and $168 \times 168 \times 244$, for PET, with a voxel size of $4.1 \times 4.1 \times 4mm^3$. CT data was acquired with the patients holding their breath, while PET was acquired with normal tidal breathing. As these are whole-body scans and we are only interested in the region of the liver, all volumes were manually cropped to include only image slices in the region of the liver.

Five clinical PET/CT data sets were used in our evaluation. For each data set the PET volumes before and after attenuation correction were registered to the CT using the registration method described above. For each data set we obtain three vector fields, two fields describe the mapping from CT coordinates to corresponding PET coordinates, and the third field is the difference between the two mappings. Descriptive statistics were computed on the magnitudes of each of the three fields. The original PET volumes were then deformed and blended with the CT volume to allow for visual assessment of the registration results.

3. RESULTS

Table 1 summarizes the results of our five experiments. In all cases the measures of location (mean, median, max) for the attenuation corrected PET are smaller than for the PET data prior to attenuation correction. That is, the magnitude of the deformations for the PET data after CT-based attenuation correction is consistently smaller than the magnitude of the deformation without attenuation correction. This is not surprising as attenuation correction is based on the CT data, which is expected to bias the registration. In all cases the location measures for the magnitude of the difference vectors are greater than that of the attenuation corrected PET data, suggesting that the deformation obtained when using the PET data without attenuation correction differs considerably from that obtained when using attenuation corrected PET. As both PET data sets are spatially equivalent and a ground truth transformation is not available, it is not clear which deformation is closer to the true one, and hence which data set should be used for registration.

Data Set	Vector Field	mean	median	std	max
1	PET with attenuation	1.9	1.8	1.0	4.9
	PET without attenuation	2.7	2.7	1.2	5.7
	difference	2.5	2.3	1.4	7.9
2	PET with attenuation	1.6	1.1	1.2	4.8
	PET without attenuation	2.7	2.7	1.1	6.0
	difference	3.1	3.0	1.4	6.6
3	PET with attenuation	2.8	2.9	1.2	5.5
	PET without attenuation	4.6	5.0	1.6	7.0
	difference	6.6	6.8	2.2	10.7
4	PET with attenuation	0.0	0.0	0.0	0.0
	PET without attenuation	4.3	4.4	1.6	6.8
	difference	4.3	4.4	1.6	6.8
5	PET with attenuation	2.8	2.7	1.3	5.7
	PET without attenuation	3.2	3.3	1.4	6.8
	difference	3.2	3.0	1.7	8.4

Table 1. Descriptive statistics of the magnitude of the vector fields that align the PET data to the CT. All measurements are in mm.

Lacking a ground truth transformation, a surrogate figure of merit for registration is often defined using manually marked pairs of homologous points on PET and CT. After registration, the points on CT are mapped to the PET coordinate system using the computed transformation and the distances between the expert delineated points and the mapped points serve as the figure of merit for the registration results. As this approach relies on visual assessment of both PET and CT data sets by experts it is not directly applicable in our case as we target structures that are visible in PET and not on CT. We thus use visual inspection only as a qualitative measure, inspecting fused PET/CT data before and after registration, as presented in Figure 2 and Figure 3, for the first and third data sets from Table 1. Unfortunately, in all cases visual inspection did not produce a distinction between registration using PET before attenuation correction and after attenuation correction. In both cases the registration seems to improve the alignment in a similar manner, though the deformation fields differ.

4. DISCUSSION AND CONCLUSIONS

We are currently developing a PET/CT based interventional navigation system to enable biopsies and RFA treatments in tumors that are clearly visible on PET and not on CT. A prerequisite of our system is accurate spatial alignment of both data sets. In the diagnostic setting, researchers have shown that the alignment of PET/CT data sets can be improved using non-rigid registration.[10] This improvement is quantified using homologous points that are visually identified across data sets.

In our institute data acquisition results in a CT data set and two PET data sets, with and without CT-based attenuation correction. The goal of this study was to evaluate the effect of the attenuation correction on non-rigid registration. As both PET data sets correspond to the same physical situation the deformations computed by the registration algorithm should be similar. As a ground truth transformation is not available we used visual inspection as an indicator of success.

While visual inspection only provides a qualitative measure, quantitative, physically meaningful, evaluation using clinical data involves developing finite element models. This approach has been previously used to evaluate the FFD based non-rigid registration with MR images of the breast.[13] As we do not have the material properties for the various soft tissue structures in the region of the liver we could not apply this approach. As a quantitative measure we only use the vector magnitudes of each of the two deformation fields.

From the quantitative analysis of the magnitude of the deformations, it is clear that CT-based attenuation correction causes bias in the registration, as the magnitude of the deformation without attenuation correction is higher than with it. This is not surprising, as CT-based attenuation correction modifies the PET intensity values in a manner that increases the correlation between the PET and CT data. Previous studies have not
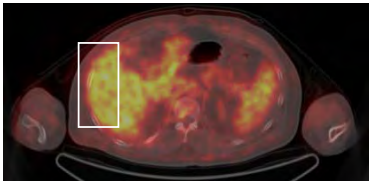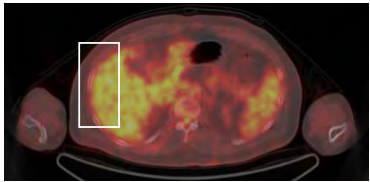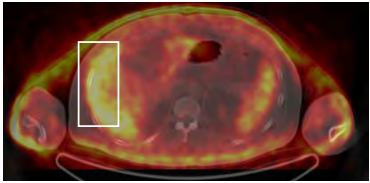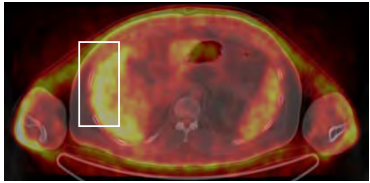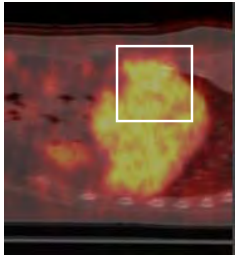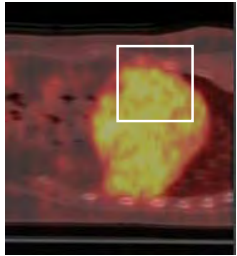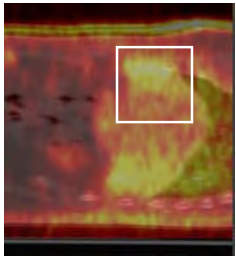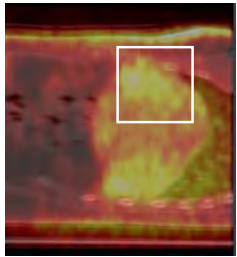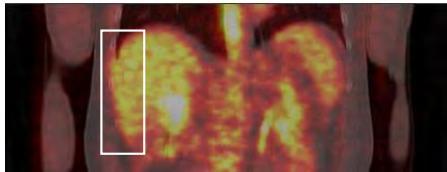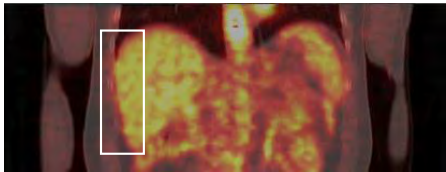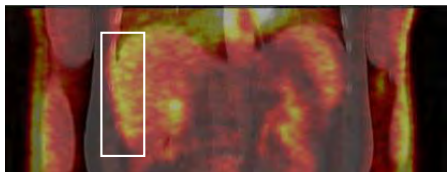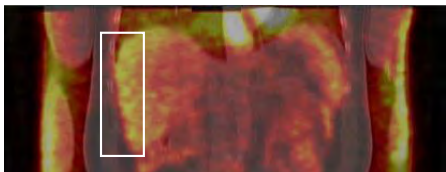
Blended PET/CT	Before Registration	After Registration
PET with CT-based attenuation correction		
PET without CT-based attenuation correction		
PET with CT-based attenuation correction		
PET without CT-based attenuation correction		
PET with CT-based attenuation correction		
PET without CT-based attenuation correction		

Figure 2. Image overlay of PET and CT, displaying the same axial, sagittal, and coronal slices with and without attenuation correction, before and after registration. Images are from the first data set in Table 1. Color figure appears in electronic version.
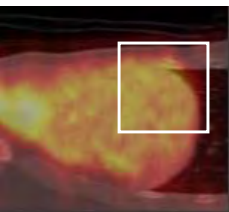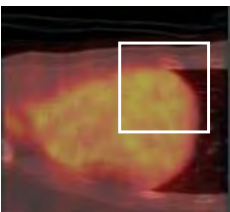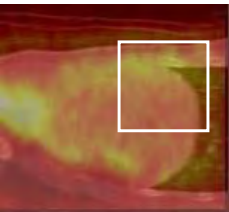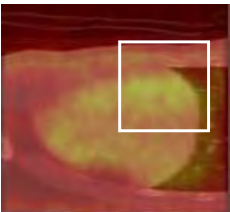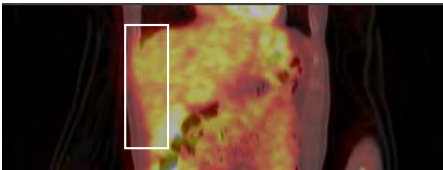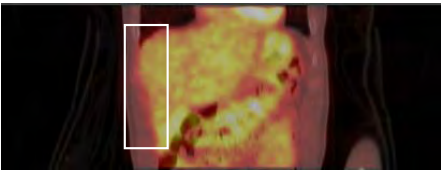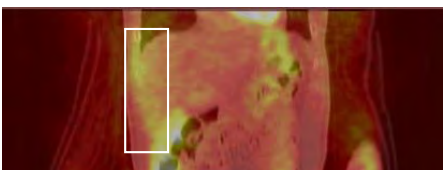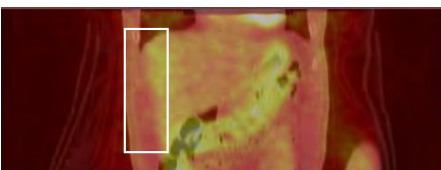
Blended PET/CT	Before Registration	After Registration
PET with CT-based attenuation correction		
PET without CT-based attenuation correction		
PET with CT-based attenuation correction		
PET without CT-based attenuation correction		
PET with CT-based attenuation correction		
PET without CT-based attenuation correction		

Figure 3. Image overlay of PET and CT, displaying the same axial, sagittal, and coronal slices with and without attenuation correction, before and after registration. Images are from the third data set in Table 1. Color figure appears in electronic version.
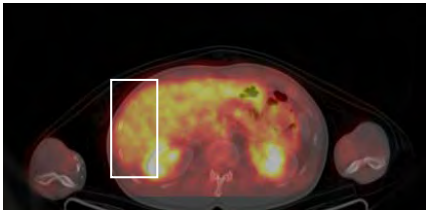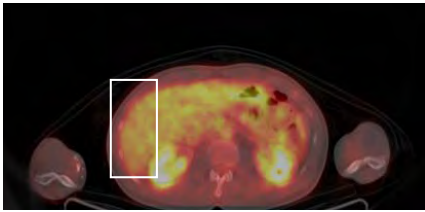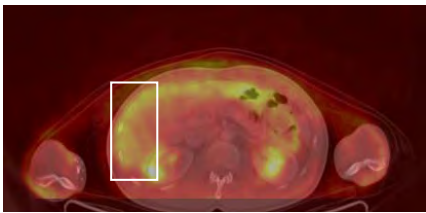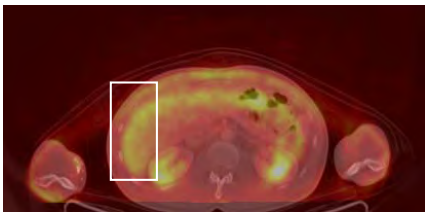
made this distinction, and in all likelihood they have used the attenuation corrected data which is usually the data set used for diagnostic purposes.

The interesting observation is that the difference in magnitude cannot be detected visually, as in all cases registration was judged as successful. This brings into question the use of PET/CT non-rigid registration as a method for accurately localizing regions of high metabolic activity on CT data. We have shown that the registration is dependent on the attenuation correction, and that visual inspection cannot distinguish between differing registration results, as it is only affected by object overlap and not by the variations within the objects. More importantly, this casts doubt on the use of visually identified homologous points as a quantitative measure of success, as these are usually identified on the surface of anatomical structures and not within them.

Our results concur with a recently published simulation study comparing the performance of multiple non-rigid registration algorithms.[14] This study has shown that widely differing deformations will result in similar overlaps which cannot be distinguished visually as the deformation differences are within the structures.

We conclude that the accuracy required by our system cannot be obtained using non-rigid registration of the PET/CT data sets currently available in our institute. To pursue the development of our navigation system we will need to interact with the imaging apparatus with the goal of acquiring 4D PET/CT data, as this will ensure that the PET and CT are spatially aligned.

ACKNOWLEDGMENTS

REFERENCES

1. K. H. Wong, E. Levy, Z. Yaniv, F. Banovac, D. Earl-Graef, and K. Cleary, "Integrated PET/CT guidance system for oncologic interventional radiology," in *Nuclear Science Symposium, Medical Imaging Conference*, 2006.
2. T. F. Hany, H. C. Steinert, G. W. Goerres, A. Buck, and G. K. von Schulthess, "PET Diagnostic Accuracy: Improvement with In-Line PET-CT System: Initial Results," *Radiology* **225**(2), pp. 575–581, 2002.
3. P. E. Kinahan, B. H. Hasegawa, and T. Beyer, "X-ray-based attenuation correction for positron emission tomography/computed tomography scanners," *Seminars in Nuclear Medicine* **33**(3), pp. 166–179, 2003.
4. P. E. Kinahan, D. W. Townsend, T. Beyer, and D. Sashin, "Attenuation correction for a combined 3D PET/CT scanner," *Med. Phys.* **25**(10), pp. 2046–2053, 1998.
5. S. A. Nehmeh *et al.*, "Quantitation of respiratory motion during 4D-PET/CT acquisition," *Med. Phys.* **31**(6), pp. 1333–1338, 2004.
6. J. W. H. Wolthaus *et al.*, "Fusion of respiration-correlated PET and CT scans: correlated lung tumour motion in anatomical and functional scans," *Phys. Med. Biol.* **50**, pp. 1569–1583, 2005.
7. M. Dawood, N. Lang, X. Jiang, and K. P. Schäfers, "Lung motion correction on respiratory gated 3-D PET/CT images," *IEEE Trans. Med. Imag.* **25**(4), pp. 476–485, 2006.
8. V. Walimbe and R. Shekhar, "Automatic elastic image registration by interpolation of 3D rotations and translations from discrete rigid-body transformations," *Medical Image Analysis* **10**(6), pp. 899–914, 2006.
9. D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "PET-CT image registration in the chest using free-form deformations," *IEEE Trans. Med. Imag.* **22**(1), pp. 120–128, 2003.
10. R. Shekhar, V. Walimbe, S. Raja, V. Zagrodsky, M. Kanvinde, G. Wu, and B. Bybel, "Automated 3-dimensional elastic registration of whole-body PET and CT from separate or combined scanners," *Journal of Nuclear Medicine* **46**(9), pp. 1488–1496, 2005.
11. L. Ibáñez, W. Schroeder, *et al.*, *The ITK Software Guide*, Kitware Inc., 2005.
12. D. J. Hawkes *et al.*, "Tissue deformation and shape models in image-guided interventions: A discussion paper," *Medical Image Analysis* **8**(2), pp. 163–175, 2005.
13. J. A. Schnabel *et al.*, "Validation of non-rigid image registration using finite element methods: Application to breast MR images," *IEEE Trans. Med. Imag.* **22**(2), pp. 238–247, 2003.
14. T. Rohlfing, "Transformation model and constraints cause bias in statistics on deformation fields.," in *Medical Image Computing and Computer-Assisted Intervention*, pp. 207–214, 2006.

8.9 Zhang 2006a: 3D Freehand Ultrasound Calibration …

Reprint begins on the next page and is nine pages.

3D Freehand Ultrasound Calibration
using an Electromagnetically Tracked Needle

Hui Zhang[a], Filip Banovac[a,b], Amy White[b], Kevin Cleary[a]
[a]Imaging Science and Information Systems ISIS) Center,
Department of Radiology, Georgetown University Medical Center
2115 Wisconsin Ave. NW, Suite 603, Washington, DC, 20007, USA
[b]Department of Radiology, Georgetown University Hospital
3800 Reservoir Road, Washington, DC, 20007, USA

ABSTRACT

3D freehand ultrasound allows intra-operative imaging of volumes of interest in a fast, cost-effective, and flexible way. However, the ultrasound device must be calibrated before it can be registered with other imaging modalities. We present a needle-based electromagnetic approach for calibrating 3D freehand ultrasound as a prerequisite for creating an intra-operative navigation system. Although most existing calibration methods require a complex and tedious experiment using a customized calibration phantom, our method does not. The calibration set-up requires only a container of water and only several frames (three to nine) to detect an electromagnetically tracked needle tip in a 2D ultrasound image. The tracked needle is dipped into the water and moved freehand to locate the tip in the ultrasound imaging plane. The images that show the needle tip are recorded and the coordinates are manually or automatically identified. For each frame, the pixel indices, as well as the discrete coordinates of the tracker and the needle, are used as the inputs, and the calibration matrix is reconstructed. Three group positions, each with nine frames, are recorded for calibration and validation. Despite the lower accuracy of the electromagnetic tracking device compared to optical tracking devices, the maximum RMS error for calibration is 1.22mm with six or more frames, which shows that our proposed approach is accurate and feasible.

Keywords: 3D freehand ultrasound, ultrasound calibration, electromagnetic tracking

1. INTRODUCTION

Three-dimensional freehand ultrasound has great promise for many surgical applications, especially in minimally invasive procedures. Potential applications include 3D anatomic visualization of pathologic tissue and various ultrasound-guided procedures. Freehand tracked ultrasound can been used to create images for organ volume measurements. We are currently developing a surgical navigation system for abdominal procedures based on the electromagnetic tracking device. We have successfully completed several phantom and pig studies, but the system has yet to incorporate real-time validation of anatomical locations and respiration deformation. For this reason, we introduced a 3D freehand ultrasound device into our navigation system. The advantages of the ultrasound-based system are that it provides real-time images, is cost-effective, and can be easily integrated into the operating room.

A requirement for 3D freehand ultrasound imaging is a motion tracking system that monitors the position and orientation of a tracked transducer. Other investigators have created 3D freehand ultrasound with tracking systems that utilize electromagnetic sensors, optical sensors, acoustic speak gaps, or a combination of these types [1-8]. The advantage of a 3D freehand tracked ultrasound system over a standard, commercial ultrasound machine with 3D image reconstruction capabilities is the ability to acquire ultrasound images in a universal coordinate frame, to which other tracked surgical instruments can be attached. In addition, the movement of the tracked probe is not restricted by a small range or limited direction. By tracking the device in a common coordinate system with the patient, the surgeon can be provided with 3D information registered to the patient anatomy.

The first step required to integrate ultrasound into these systems is a calibration step as presented in this paper. In ultrasound image calibration, the 3D physical dimensions of the beam plane are computed with respect to the tracked device. After the calibration process is completed, standard 2D B-mode ultrasound images can be reconstructed into a 3D volume using the positional information obtained from the tracking system.

There are four general methods to construct an ultrasound volume: 1. constrained sweeping techniques; 2. 3D probes; 3. sensorless techniques; and 4. 2D tracked probe (freehand) techniques [9]. Calibration must be performed to determine the position of the ultrasound coordinate system in relation to the reference coordinates of the tracking system. This calibration is most often performed by registration between the ultrasound image and a geometrical model of the phantom reconstructed from a CT image, or by touching a pre-calibrated pointer. These methods are referred to as point-based calibration methods, and are implemented by matching corresponding points from two different coordinate frames. The high accuracy of point-based calibration methods has been demonstrated by several research groups [1,10]. These techniques carry the advantage of having a simple experiment setup, but can be more time consuming than other automated techniques, such as surface- and intensity-based mappings, because the identification of the points in the images is a manual process.

The results of several ultrasound beam calibration techniques indicate that the point-based "cross-wire" phantom produced the most precise and accurate reconstructions. Overall, an optical tracking system will provide a better calibration results than electromagnetic systems since optical systems are more accurate and not subject to disturbance from metal objects. In the "cross-wire" phantom study, a mean precision of 1.65mm was obtained using an optical tracking system, while an RMS error of 2.41mm was obtained using a electromagnetic tracking system [11]. Instead of using specific calibration phantoms, we use a tracked needle as the indicator to avoid the registration and touching steps that can introduce unnecessary calibration errors. Several researchers have experimented with different approaches using point objects (such as a pinhead or a small suspended head) instead of the cross-wire phantom. For example, Muratore and Galloway used the tip of an optically tracked pointer to find the spatial transformation [12]. We use an electromagnetically tracked needle, and take a different mathematical approach to calculate the calibration matrix.

2. MATERIALS AND METHODS

2.1. Tracked ultrasound image acquisition

The ultrasound device used in the experiment is a portable system from Terason Inc. (put city, state here). It has 128 channels and connects with the computer through a standard IEEE 1394 firewire port. The electromagnetic tracking device is the AURORA from Northern Digital Inc (Ontario, Canada). A frame grabber (vendor?) is used to obtain the ultrasound image. We have found the tracking accuracy of this system to be minimally affected by small metal instruments in close proximity. The tracker attached to the ultrasound probe must be a 6 degree of freedom (DOF) tracker, which includes translation and rotation. The general set-up is shown in Figure 1. The tracked ultrasound probe is fixed by a plastic holder, and the bean surface contacts the water. The tracked needle is placed in a plastic cube in the water, which stabilizes the needle and can reduce the error generated by small unintentional hand movements.

The frame grabbed ultrasound image is shown as Figure 2. The bright line at the bottom of the image is the water container. Because the entire experiment is set-up in a container of water, the contrast allows adequate visualization of the needle, and the needle tip can therefore be easily identified manually. On ultrasound the needle tip is not sharply demarcated, but is located at the center of the bright signal observed on imaging as shown in Figure 2.

Fig.1 Experimental setup (electromagnetic field generator,
tracked ultrasound probe, needle in plastic block, water container)
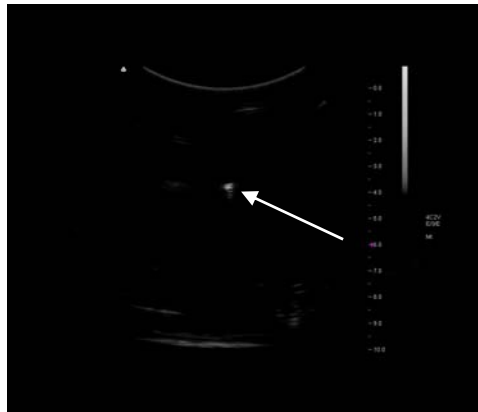
Fig.2 Ultrasound image (with needle tip at bright spot indicated by arrow)

2.2. Pivot calibration of electromagnetic tracker

Before performing the calibration, we pivoted the needle to get the correct offset from the tracked sensor position to the tip [13]. By applying the calibration transforms into the tracking information, the registration step of physical and tracking space, discussed previously in the point-based calibration methods, will be avoided. The position and orientation information reported by the AURORA tracking system is the center position of the sensor coil. When the sensor coil is embedded in a surgical tool, the transformation between the coil and the tip of the tool is required. This information is typically determined by a pivot calibration, in which the tip of the instrument is placed in a small holes, or divot, and the instrument is rotated back and forth as sensor data is collected.

The AURORA system returns the translational position of the sensor coil relative to a coordinate system fixed at the center of the field generator. The orientation of the sensor is also returned and specified using quaternions. Typical AURORA sensors are five degree of freedom sensors in that they cannot resolve their orientation about their long axis (the roll axis). The tracked needle used in this experiment is 5DOF. From the translational and orientation information, a 4 by 4 homogenous transformation matrix can be constructed. When the sensor coil is embedded into the needle, the long axis of the coil is parallel to the needle such that the only offset from the coil to the needle tip is along the z-axis of the needle. The transformation from the field generator coordinate system to the needle tip coordinate system can then be calculated from:

$$\begin{bmatrix} r_{00} & r_{01} & r_{02} & t_x \\ r_{10} & r_{11} & r_{12} & t_y \\ r_{20} & r_{21} & r_{22} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ offset \\ 1 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix}$$

(1)

In this equation, R is the rotation matrix, T is the translation vector and (x_0, y_0, z_0) is the pivot position (in this case the tip of the needle).

Typically, we record several hundred samples while pivoting the needle. Equation (1) can be re-written as follows where the unknown variables of offset and (x_0, y_0, z_0) can be seen:

$$r_{02} \cdot offset - 1 \cdot x_0 + 0 \cdot y_0 + 0 \cdot z_0 = -t_x$$
$$r_{12} \cdot offset + 0 \cdot x_0 - 1 \cdot y_0 + 0 \cdot z_0 = -t_y$$
$$r_{22} \cdot offset + 0 \cdot x_0 + 0 \cdot y_0 - 1 \cdot z_0 = -t_z$$

(2)

These equations can then be written as:

$$A * offset + B * x_0 + C * y_0 + D * z_0 = E$$

(3)

In equation (3), A, B, C, D and E are the column coefficient vectors, which can then be re-written as:

$$M \cdot \begin{bmatrix} offset \\ x_0 \\ y_0 \\ z_0 \end{bmatrix} = N$$

(4)

Since M is not a square matrix, the unknowns can be found using the singular value decomposition (SVD) or Moore-Penrose inverse:

$$\begin{bmatrix} offset \\ x_0 \\ y_0 \\ z_0 \end{bmatrix} = \left(M^T \cdot M\right)^{-1} \cdot M^T \cdot N$$

(5)

Additionally, the RMS error can be computed as:

$$rms = \sqrt{\left| M \cdot [offset \quad x_0 \quad y_0 \quad z_0]^T - N \right|^2 / num}$$

(6)

where *num* is the number of samples.

2.3. Calibration of the ultrasound probe

To set up the experiment, we placed the ultrasound probe into a container of water. After the probe and the field generator were fixed at several different positions, the tracked needle was dipped into the water and moved to determine the exact position where the tip intersected with the imaging plane. The transformation relationship of the calibration system is shown in Figure 3.
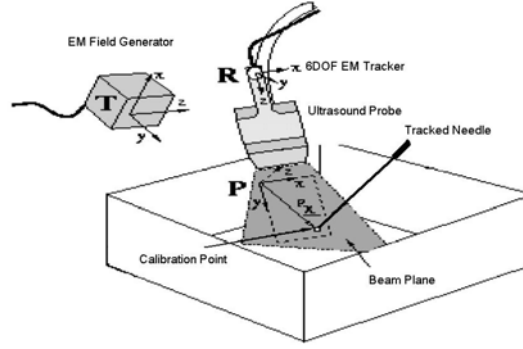
Fig.3 Calibration coordinates system

The position of the needle tip, P_{tip}, can be determined by its pixel location in the ultrasound image, P_{us}, the calibration matrix, T_c, and the reference tracker's transformation matrix, T_{ref}, using the following equation:

$$P_{tip} = T_{ref} \cdot T_c \cdot P_{us} \tag{7}$$

where $P_{us} = (u, v, 0, 1)^T$, and u and v are the column and row indices of the pixel in the grabbed image. The index along the z-axis is set to 0 because we used a 2D beam. The scale factors are integrated with the calibration matrix, T_c, such that the extra scale factor in the computation is not required. After left multiplying the inverse of the reference transformation matrix T_{ref}, equation (7) can be written as:

$$P_{tip/ref} = T_{ref}^{-1} \cdot P_{tip} = T_c \cdot P_{us} = T_c \cdot (u, v, 0, 1)^T \tag{8}$$

where $P_{tip/ref}$ is a vector. After taking n frames, equation (8) is accumulated as:

$$\left(P_{tip/ref,1}, \cdots P_{tip/ref,n} \right) = T_c \cdot \begin{pmatrix} u_1 & & u_n \\ v_1 & \cdots & v_n \\ 0 & & 0 \\ 1 & & 1 \end{pmatrix} = T_c \cdot T_{us} \tag{9}$$

where $T_{us} = \begin{pmatrix} u_1 & & u_n \\ v_1 & \cdots & v_n \\ 0 & & 0 \\ 1 & & 1 \end{pmatrix}$. The calibration matrix is calculated by SVD solution using

$$T_c = \left(P_{tip/ref,1}, \cdots P_{tip/ref,n} \right) \cdot T_{us}^T \cdot \left(T_{us} \cdot T_{us}^T \right)^{-1} \tag{10}$$

This computation results in a matrix with a zero vector in the third column, for the T_{us} contains a zero vector and all 1 vector in the third and forth rows. The non-zero full matrix is computed with Horn's method [14].

2.4. Algorithm validation

The validation of the calibration is significant since it confirms that the calibration matrix computed will accurately reconstruct the 3D plane in the tracking space. There are four general methods to validate the calibration matrix: the first is to calculate the RMS of the solution; the second is to use some other points which are not included in calibration computation to verify the transformation accuracy; the third is to measure the orthgonality and normality conditions of the rotation matrix of the calibration term; and the fourth is to compare the resliced image with the grabbed ultrasound image plane directly.

The general transformation matrix by the rotation and translation is

$$T = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{11}$$

Among them, r11-r33 are the elements of rotation matrix, and should obey the orthogonality and normality conditions. t_x, t_y and t_z are the translation elements. The indices vector with the scalers is

$$V = \begin{bmatrix} s_x \cdot u \\ s_y \cdot v \\ s_z \cdot 0 \\ 1 \end{bmatrix} \tag{12}$$

s_x and s_y are the scale factors and are changed according to the depth and some other parameters from ultrasound device. After the matrix and the vector are multiplied together, we get

$$T \cdot V = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} s_x \cdot u \\ s_y \cdot v \\ s_z \cdot 0 \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} \cdot s_x & r_{12} \cdot s_y & r_{13} \cdot s_z & t_x \\ r_{21} \cdot s_x & r_{22} \cdot s_y & r_{23} \cdot s_z & t_y \\ r_{31} \cdot s_x & r_{32} \cdot s_y & r_{33} \cdot s_z & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 0 \\ 1 \end{bmatrix} \tag{13}$$

The SVD method we used in the calibration is to calculate the final matrix:

$$T_c = \begin{bmatrix} r_{11} \cdot s_x & r_{12} \cdot s_y & r_{13} \cdot s_z & t_x \\ r_{21} \cdot s_x & r_{22} \cdot s_y & r_{23} \cdot s_z & t_y \\ r_{31} \cdot s_x & r_{32} \cdot s_y & r_{33} \cdot s_z & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{14}$$

At this time, only the first two columns scalar s_x and s_y are compared for the third column is not meaningful for the 2D ultrasound probe. The orthogonality should be 0.0 in the idea condition.

3. RESULTS

In the experiment, we recorded all the data for three different group positions. For each group, nine different indices of the needle tip and the tracking information were recorded as the input. In each group, three to nine frames were used to compute the calibration matrix using the previously discussed equations. Two groups served as target point sets for validating the generated matrix. The results are shown in Table 1. Although the minimum requirement to calculate the calibration matrix is three frames, six or more frames provide increased accuracy. In these data, the maximum RMS calibration error is 1.22mm and the maximum target validation error is 3.45mm. Considering that we used the electromagnetic tracking system, which is not as accurate as an optical tracking system, our proposed method seems reasonably accurate. For each group, the first row is the RMS calibration error with different frame number, and the following two rows are the respective RMS target validation error. The bold value is the maximum error using six or more points for calibration.

Calibration Group	Validation Group	Number of points						
		3	4	5	6	7	8	9
1		0.00	0.47	0.55	0.60	0.70	**1.22**	1.19
1	2	3.13	2.35	2.40	2.42	1.74	1.58	1.78
1	3	3.28	2.63	2.81	2.79	2.43	3.21	**3.45**
2		0.00	0.48	0.59	0.60	0.57	0.66	0.76

2	1	4.77	4.55	3.07	2.50	2.62	2.45	2.20
2	3	5.95	4.63	1.74	2.09	2.01	1.73	2.06
3		0.00	0.42	0.49	0.46	0.44	0.45	0.44
3	1	15.43	3.04	3.22	3.28	3.33	3.16	3.21
3	2	8.31	2.47	2.34	2.37	2.32	2.32	2.32

Table 1 Calibration and validation result (mm)

The validation method we proposed is also verified by three different groups. For each group, the orthogonality and the s_x and s_y are calculated respectively. The orthogonality should be close to 0.0 and the s_x and s_y should remain constant under ideal conditions. Through the validation result, as shown in Table 2, our calibration method gives a reasonable output and is considered to fit those orthogonality and scale conditions. When the frame number is increased to more than six, the orthogonality is close to 0.0 and the scale of the X and Y axis are consistent.

(a) Group 1			
Frame	Orthogonality	Scale X	Scale Y
3	1.037e-3	0.160	0.156
4	1.346e-3	0.158	0.156
5	1.520e-3	0.158	0.155
6	1.302e-3	0.158	0.157
7	6.481e-4	0.155	0.157
8	5.286e-5	0.153	0.158
9	-3.4278e-4	0.152	0.157

(b) Group 2			
Frame	Orthogonality	Scale X	Scale Y
3	-6.413e-3	0.147	0.173
4	-5.482e-3	0.151	0.169
5	-2.292e-4	0.152	0.170
6	2.668e-4	0.153	0.161
7	-5.568e-5	0.153	0.157
8	-5.871e-5	0.153	0.157
9	-5.919e-5	0.153	0.156

(c) Group 3			
Frame	Orthogonality	Scale X	Scale Y
3	-1.031e-2	0.156	0.131
4	-1.097e-3	0.156	0.161
5	-6.272e-4	0.155	0.157
6	-8.752e-4	0.155	0.157
7	-7.873e-4	0.155	0.156
8	-5.489e-4	0.156	0.157
9	-4.889e-4	0.156	0.156

Table 2 Validation of orthogonality and scale of calibration matrix

After the calibration, the output matrix is used in our image-guided surgical system, Navigator, to display the freehand ultrasound image plan. The reconstructed 2D beam plane is displayed in 3D to show the correct relationship with the volume-rendered image. The reconstructed CT image can also be displayed simultaneously to compare the difference between the two imaging modalities, providing more information for the planned intervention. The results from the phantom study are shown in Figure 4. Several skin fiducials and one tracked needle are used for registration. After calibration and registration, the related ultrasound image and reformatted CT image are displayed together, in which the simulated aorta is identified. The co-registered images from a swine study are also shown in Figure 5. The swine study was part of an approved protocol. As a part of the open source project titled Image Guided Surgical Toolkit (IGSTK), the source code of the pivot calibration and point-based calibration parts can be downloaded from www.igstk.org.

4. CONCLUSIONS

We have performed ultrasound calibration using the tip of an electromagnetically tracked needle as the calibration device, rather than a phantom. A new mathematical linear algebra matrix operation was derived to solve the calibration matrix. The validation method of our algorithm is also included. The results of our experiments prove that this is a feasible calibration method for tracking-based surgical navigation systems.

There are a number of requirements for proper implementation of our point-based calibration method. First, a tracked sensor must be attached to the ultrasound probe, which is necessary for image-guided procedures. Second, a good image, in which the tip of the tracked needle is easily identified, is also required. This can be seen from the grabbed image shown in Figure 2. If we fix the ultrasound probe in a stable position and move the tracked needle, the exact position of where the needle tip is passing by the beam plane can be easily identified by image subtraction. The tracker information with the same time stamp will be recorded and set as the input as well as the pointer coordinate. This will greatly accelerate the calibration speed and requires only minutes for accurate calibration.

Fig. 4 Phantom setup and the US-CT image display in Navigator system

Fig. 5 US-CT image display from swine study showing the swine gallbladder

REFERENCES

1. R. W. Prager, R. N. Rouhling, A. H. Gee, and L. Berman, Rapid calibration for 3-D freehand ultrasound, Ultrasound in Med. & Biol., 24(6):855-869, 1998
2. Pagoulatos N, Haynor DR, Kim Y. A fast calibration method for 3-D tracking of ultrasound images using a spatial localizer. Ultrasound Med Biol. 27(9):1219-1229, 2001
3. Nicolas Andreff, Radu Horand, Bernard Espiau, Robot Hand-Eye Calibration Using Structure-from-Motion, The International Journal of Robotics Research, 20(3):228-248, 2001
4. A. Viswanathan, E.M. Boctor, R.H.Taylor, G.D. Hager and G. Fichtinger. Immediate ultrasound calibration with three poses and minimal image processing. MICCAI 2004, Saint-Malo, France. 3217:446-454, 2004
5. Boctor, Emad M.; Jain, Ameet; Choti, Michael A.; Taylor, Russell H.; Fichtinger, Gabor, Rapid calibration method for registration and 3D tracking of ultrasound images using spatial localizer, Medical Imaging 2003: Ultrasonic Imaging and Signal Processing. Edited by Walker, William F.; Insana, Michael F. Proceedings of the SPIE, Volume 5035:521-532, 2003
6. Sebastian Eulenstein and Thomas Lange and M. Hünerbein and Peter Schlag and Hans Lamecker, Ultrasound Based Navigation System Incorporating Preoperative Planning for Liver Surgery, in: Proc. Computer Assisted Radiology and Surgery, 2004
7. Lionel G Bouchet, Sanford L Meeks, Gordon Goodchild, Francis J Bova, John M Buatti, and William A Friedman, Calibration of three-dimensional ultrasound images for image-guided radiation therapy, Phys. Med. Biol. 46:559-577, 2001
8. Aaron Fenster, Donal B Downey, and H Neale Cardinal, Three-dimensional ultrasound imaging, Phys. Med. Biol. 46:67-99, 2001
9. Laurence Mercier, Thomas Lango, Frank Lindseth, D. Louis Collins, A review of calibration techniques for freehand 3-D ultrasound systems, Ultrasound in Med. & Biol., 31(4):449-471, 2005
10. Blackall JM, Rueckert D, Maurer CR, er al. An image registration approach to automated calibration for freehand 3D ultrasound. Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2000, 1935:462-471, 2000
11. Leotta DF, Detmer PR, Martin RW. Performance of a miniature magnetic position sensor for three-dimensional ultrasound imaging. Ultrasound in Med. & Biol., 23(4):597-609, 1997
12. Diane M. Muratore, Robert L. Galloway, Beam calibration without a phantom for creating a 3-D free handultrasound system, Ultrasound in Med. & Biol., 27(11):1557-1566, 2001
13. Kevin Cleary, Hui Zhang, Neil Glossop, Elliot Levy, Filip Bonavac, Electromagnetic Tracking for Image-Guided Abdominal Procedures: Overall System and Technical Issues, IEEE EMBC 2005, Shanghai, China, Sep 1-4, 2005
14. Berthold K. P. Horn, Closed-form Solution of Absolute Orientation using Unit Quaternions, Journal of the Optical Society of America A, 4:629-642, 1987

8.10 Zhang 2006b: High Quality GPU Rendering …

Reprint begins on the next page and is eight pages.

High Quality GPU Rendering with Displaced Pixel Shading

Hui Zhang, Jae Choi

Imaging Science and Information Systems Center, Department of Radiology, Georgetown University

2115 Wisconsin Ave. NW, Suite 603, 20007, Washington, DC, USA

ABSTRACT

Direct volume rendering via consumer PC hardware has become an efficient tool for volume visualization. In particular, the volumetric ray casting, the most frequently used volume rendering technique, can be implemented by the shading language integrated with graphical processing units (GPU). However, to produce high-quality images offered by GPU-based volume rendering, a higher sampling rate is usually required. In this paper, we present an algorithm to generate high quality images with a small number of slices by utilizing displaced pixel shading technique. Instead of sampling points along a ray with the regular interval, the actual surface location is calculated by the linear interpolation between the outer and inner points, and this location is used as the displaced pixel for the iso-surface illumination. Multi-pass and early Z-culling techniques are applied to improve the rendering speed. The first pass simply locates and stores the exact surface depth of each ray using a few pixel instructions; then, the second pass uses instructions to shade the surface at the previous position. A new 3D edge detector from our previous research is integrated to provide more realistic rendering results compared with the widely used gradient normal estimator. To implement our algorithm, we have made a program named DirectView based on DirectX 9.0c and Microsoft High Level Shading Language (HLSL) for volume rendering. We tested two data sets and discovered that our algorithm can generate smoother and more accurate shading images with a small number of intermediate slices.

Keywords: volume visualization, GPU rendering, pixel shader, early Z-culling

1. INTRODUCTION

Interactive volume rendering is an important research area in the medical visualization. Various methods have been proposed to accelerate the general purpose central processing unit (CPU) to achieve an interactive rendering speed. However, the computing power of a single CPU is not enough to get the interactive speed for a large data set. To solve this problem, some dedicated CPU or graphics hardware, and even some special accelerated graphics chips, have been developed. Moreover, texture-based approaches have positioned themselves as efficient tools for the direct rendering of volumetric scalar fields on graphics workstations and general consumer class hardware [1-7]. Among them, GPU-based volume texture rendering is a popular method that has been used recently. To exploit hardware-assisted texture mapping techniques for volume rendering, an approach called "proxy geometry" is used, in which a stack of reformatted slices perpendicular to the view direction are assembled and blended to achieve the final rendering plane. The ability to leverage the embedded tri-linear interpolation hardware is the core of such acceleration techniques.

Typically, there are two kinds of GPU-based ray tracers: slice-based volume renderer [8] and ray marching [9]. Ray marching calculates the ray steps inside GPU, and does not exploit the polygon rendering ability of the hardware. Slice-based volume renderers draw a series of slices that appear along the viewing direction and are clipped to fit inside of the volume. These slices are composited within the frame buffer in order to approximate the integrals of rays from the eye through each pixel. Each pixel at the intersection of the ray and slices are sampled by GPU to retrieve the density information. To achieve high quality volume rendering, the sampling step for each ray must be small enough to increase the number of view-axis aligned slices for texture mapping and to decrease the rendering speed. Our method achieves high quality image rendering with a small number of slices.

2. MATERIALS AND METHOD

2.1. 3D texture-based volume rendering

Volume rendering via 3D textures is a well-known process. The multi slices, which are parallel to the view direction, are blended together to produce the final rendering images. The volume information is represented and mapped by the texture information of each slice and the hardware integrated alpha blending capability is used to accelerate the volume rendering. Generally, there are two rendering orders: back-to-front or front-to-back. In the back-to-front order, the blending equation is

$$C_{dst} = (1 - \alpha_{src}) \cdot C_{dst} + \alpha_{src} \cdot C_{src} \tag{1}$$

and in the front-to-back order, the blending equation is

$$C_{dst} = C_{dst} + (1 - \alpha_{dst}) \cdot \alpha_{src} \cdot C_{src}$$
$$\alpha_{dst} = \alpha_{dst} + (1 - \alpha_{dst}) \cdot \alpha_{src} \tag{2}$$

where C_{dst}, α_{dst} and C_{src}, α_{src} are the color and opacity values of the current color buffer and the destination color buffer respectively [8].

The whole multi-slice blending is shown in Fig. 1. The multi texture slices are blended together to produce the volumetric rendering image.
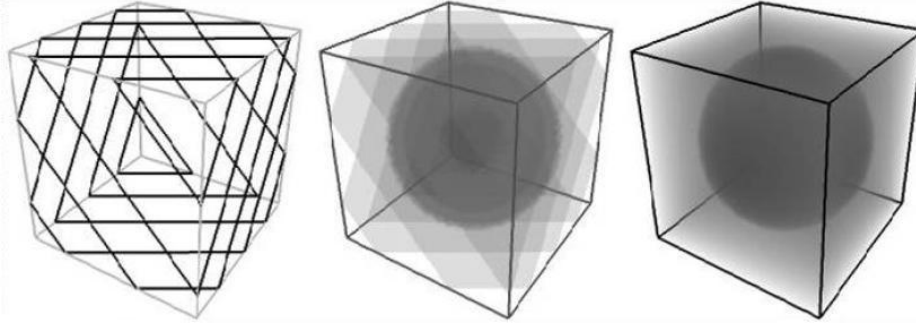

Fig. 1 Volume rendering via 3D texture [2]

Texture mapping takes a long time because of the huge amount of fragment and pixel shader operations necessary to blend the color and opacity information. Those instructions include the texture fetch, blending, and surface shading. Most of the fragment operations are executed for the intermediate buffers and will not contribute to the final image. The number of fragment and pixel operations, which determines the computation and rendering time, is generally in proportion with the number of the reformatted texture slices submitted to GPU. Thus, reducing the number of intermediate slices can accelerate the rendering; however, the quality of the final composite image will dramatically decrease when the number of the slices is reduced below a certain limit.

To solve this, we present a way to use displaced pixels to represent the iso-surface pixels instead of increasing the number of slices. For every pixel in each reformatted slice, the same amount of fragment and pixel operations are performed regardless how much it will contribute to the final rendering image or whether it is an empty voxel. In the iso-surface shading, the pre-computed normal information for each pixel is stored in the RGB channels, and the surface shading is done by pixel shader operations. Many operations are performed for the shading effect in each slice point regardless whether it is a real iso-surface point, which results in unnecessary computation. Thus, the multi-pass technique and the recently developed early Z-culling technique [10] are applied to avoid those expensive computations.

2.2. Displaced pixel shading

To employ GPU-based ray tracer to render illuminated surfaces, the alpha-testing ability is used for detecting the iso-surface. At each sample point, we fetch the corresponding value from the texture volume and test whether that the sample intensity is larger than or equal to the threshold. If the intensity is equal or greater than the iso-surface intensity, the current position of the pixel will be stored and the shading computation is performed from the pre-computed gradient channels. However, this sample position, shown as the inner point in Fig. 2, is not the exact location of the surface point. The accurate position is located between the inner point and the outer point in Fig. 2. This implies that the higher sampling rate and more view-axis slices can improve the image quality because they can give the point close to the actual surface location.
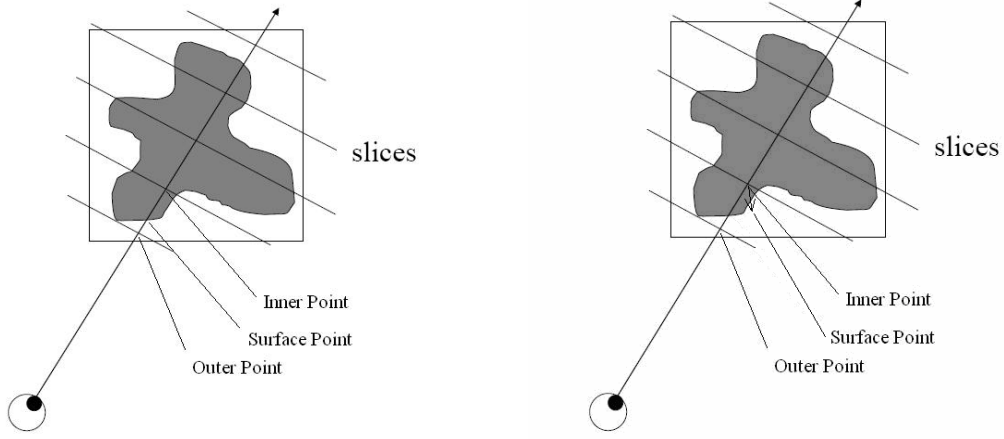
Fig. 2 Pixel displacement by view vector and normal vector

Our algorithm exploits the capacity of pixel shading language to get more accurate surface locations. After the ray reaches the first point inside the object (inner point), the previous sampling point (outer point) is retrieved. The ray vector is stored in the constant register, which is updated by the camera-view position and passed by the main program to the pixel shader. A linear interpolation is then performed between the inner point and outer point using their intensity values to get the accurate surface point position, as indicated in Fig. 2. This displaced point is used for shading computation instead of the inner sampling point. The linear interpolation function is shown in equation (3):

$$P_0 = \left(\frac{I_0 - I_{out}}{I_{in} - I_{out}} \right) \cdot P_{in} + \left(\frac{I_{in} - I_0}{I_{in} - I_{out}} \right) \cdot P_{out} \tag{3}$$

where P_{out}, P_{in}, and P_0 are the positions of the outer point, inner point, and the iso-surface point, respectively. I_{out} and I_{in} are the intensity values of each position. I_0 is the predefined iso-surface value for the alpha test. In [5], a pre-integration method was used for remove the artifacts due to the small number of slices. Their method mostly addressed on the alpha blending methods and also addressed on how to apply the pre-integration to the iso-surface shading. In their iso-surface shading algorithm, a linear interpolation of front slice's and back slice's normal was performed to increase the accuracy. That means, only the normal of the inner point and outer point are utilized to interpolate the final normal and the position of the surface point is not changed. In our method, the surface point's position is changed and then its normal is tri-linear interpolated by the voxels' normal around the final surface position, so the rendering result is different. Another extension of our method is to use the inner point's normal to displace the surface point instead of the view vector. In this way, a morphed surface point is calculated to show more details of the high-curvature positions.

The linear interpolation using the inner and outer points can be pre-calculated and stored in a 2D texture map. The x axis is set as the intensity difference between those two end points, and the y axis is the actual density at the intersection point. Thus, the instructions for linear interpolation computation can be implemented by two texture fetch instructions and one subtraction operator. For 8-bit data, the size of the 2D texture map is 512x512, which occupies only a small portion in the texture memory of the current GPUs.

2.3. Multi-pass and early Z-culling

In a volume rendering experiment [8], only 0.2% to 4% of all generated fragments contributed to the final image. Most of the fragments and pixel shader operations were wasted in the empty space and over-sampled ray path. Fig. 3 is the rendering pipeline of the DirectX. An alpha test is used in the algorithm to set the iso-surface for display. The pixel with a value lower than the alpha test value will be discarded, but the pixel shader operations for those discarded pixels will still be executed in the pipeline in front of the alpha test. Only the early Z-culling technique will prevent the unnecessary computations and stop the pixel shader operations before throwing the pixel inside the pipeline. These illumination calculations in pixel shader consume both time and resources, and should be computed only once at the surface point. Thus, a two-pass strategy, known as the early Z-culling technique, is employed in our algorithm too.

Prior to executing the pixel shader, some GPU chips check the interpolated z value against the z value in the z-buffer. This step occurs for any pixels that pass the hierarchical z test and is actually going to use the primitive's interpolated z (rather than computes z in the pixel itself). This additional check provides not only an added efficiency win when using long, costly pixel shader operations, but also a form of pixel-level control flow in specific situations. This approach is known as "early Z-culling" in real-time rendering [10]. In the iso-surface shading routine of volume rendering, a bulk of pixel shader operations are performed to illuminate the segmented iso-surface. Instead of the computationally expensive operations, we insert inexpensive rendering passes for efficiently approximating the surface.

Vertices

High-Order-Surfaces Tesselation

Source Operations

Transformation & Lighting

Vertex Shader

Face Culling
User Clip Planes
Frustrum Clipping
Homogenous Divide
Viewport Mapping

Vertex Operations

Triangle Setup
Rasterization/
Multisampling

Multitexturing DX

Pixel Shader

Fog Effects

Alpha Test

Stencil, Depth Test

Alpha Blending

Dither

Pixel Operations
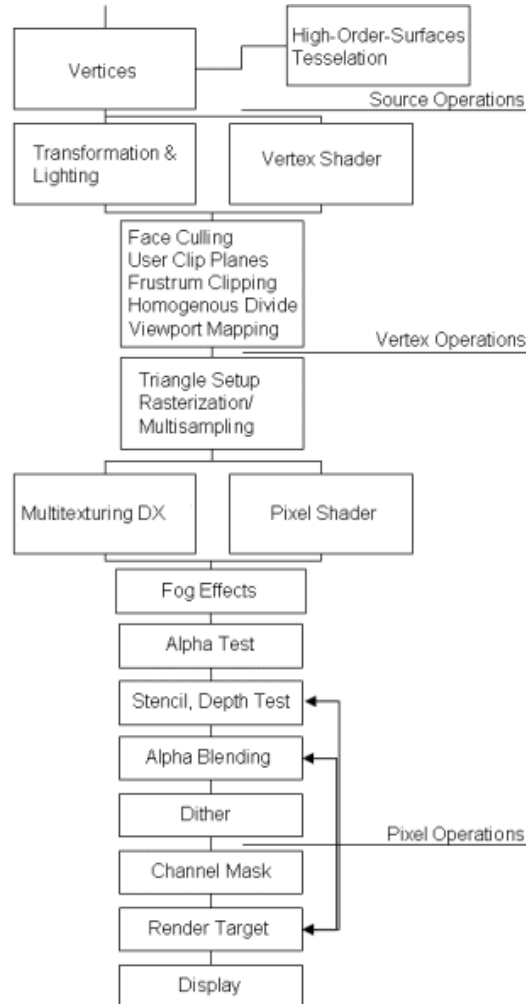
Channel Mask

Render Target

Display

Fig. 3 Rendering pipeline of DirectX, from [11]

Compared with the early Z-culling method used in [8], which used this technique to remove the non-intersection rays with the objects, we used a different early Z-culling mechanism and implementation, which not only removes the non-intersection rays, but also directly gives the surface location for the rays intersecting with the objects. In [8], a conditional code, MAX or 0, is used to indicate whether the ray tracing through this pixel will hit the object or not. If the ray won't hit the object, the following pixel shader computation along this ray will be discarded. If hit, the following passes will trace the ray using the bounding box method. In the method we proposed here, we save the depth value of the surface point instead of the conditional codes, so that in the second pass, the position of the iso-surface will be located directly, which means that not only the non-intersect rays, but also the voxels in front of the surface for each intersected ray will be discarded too. In the first pass of our rendering pipeline, only a simple volume sample fetch instruction is performed to get the depth value approximately, with the depth comparison function set to LESSEQUAL and the rendering sequence set at back-to-front. In the second pass, the depth comparison function is set to EQUAL to

locate the surface position; then, a number of pixel shading instructions are performed for illumination. The depth buffer is writable and changed in the first pass; thus, the early Z-culling is not effective at that time. However, in the second pass, the depth buffer is locked and is read-only, and the alpha test is turned off, so that the early Z-culling is performed automatically. A comparison of two methods is shown as Fig. 4. The blue part is the region where the pixel shader computation is discarded. The sphere indicates the target rendering object in the volume images, and the rectangle bounding the sphere indicates the bounding box techniques used in [8] to discard the pixel computation before the ray intersects the object.
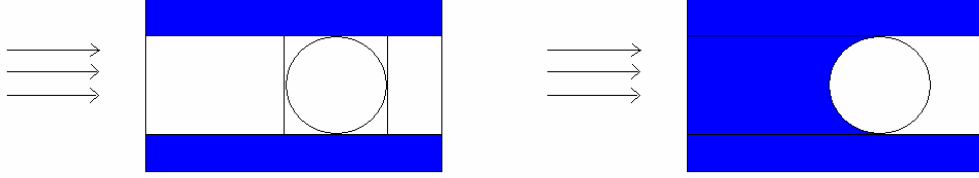

Fig. 4 Discarded pixel shader computation in
(a) J Kruger's method   (b) Our method

2.4. Moment-based 3D edge descriptor

In our shading algorithm, the simplified Phong light model is used to illuminate the surface, as follows:

$$I = I_\alpha + I_d \cdot \left(\vec{V} \cdot \vec{N} \right) \tag{4}$$

where I_α is the ambient light intensity, I_d is the diffuse light intensity, $\vec{V}$ is the direction of the view and $\vec{N}$ is the gradient normal. The first two parameters I_α and I_d are adjusted by the user, and the view direction is changed each time when the view matrix is changed. Normal gradient information is the only component that comes from the original volume dataset and reflects the real image content. The widely used gradient operator is computed by

$$G = \left(G_x = \frac{I_{x+1} - I_{x-1}}{2}, G_y = \frac{I_{y+1} - I_{y-1}}{2}, G_z = \frac{I_{z+1} - I_{z-1}}{2} \right) \tag{5}$$

It is easy to implement, but the result is not smooth and sensitive to noise because it does not accounts for diagonal components around the designated location to approximate the gradient. In our rendering algorithm, a moment-based 3D edge descriptor [12] is used to calculate a smoother and more accurate normal vector, which gives a much better shading result than the gradient normal. An orthogonal Legendre moment and a two-order sphere surface model are used together to compute the distribution of each neighbor pixel in a window with any size. The masks of the edge operator and the normal information is pre-computed and stored in the RGB channels, so that the interactive rendering speed will not be affected.

2.5. DirectView volume rendering software

Based on the accelerated algorithms presented in this paper, a volume rendering software called DirectView was developed as a platform for interactive medical image visualization. For the implementation, we have used DirectX 9.0c with Microsoft's Effect framework and HLSL. The software can open a DICOM image directory or raw volume dataset and can provide interactive volume rendering. Fig. 5 shows the screen capture of this software. This software and the manual can be found at http://www.huizhang.org/directview/.

The projection modes have alpha blending, maximum intensity projection, and sum intensity projection. The texture rendering modes inside the pipeline include 2D texture mapping, 3D texture mapping, iso-surface shading, displaced iso-surface shading, and composite shading. Front-to-back and back-to-front rendering sequences are supported for each mode. Gradient-based and moment-based normal can be selected by the user. Several pre-set transfer function definitions are stored and the transfer function can be designed on the fly. Some other accelerated techniques such as rendering-to-surface are also used to reduce the rendering area. In the programmable mode, users can design their own HLSL code to test the rendering and GPU-based image processing algorithms.
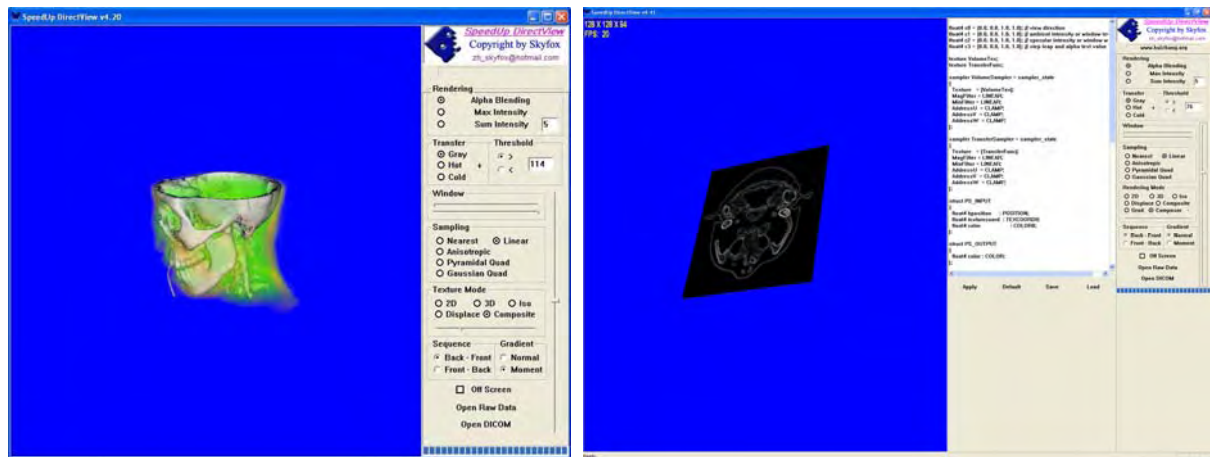
Fig. 5 DirectView software, composite rendering and edge detection

The system provides an effect editor to manually write the pixel shader fragment. The general pixel shader code for displaced pixel shading is outlined below. It has two different pixel shader fragments for each pass. The function of the first pass is to locate the iso-surface position, and only one texture fetch instruction is required. In the second pass, the pre-computed gradient normal is retrieved and illuminated, as well as the computation of the position of the displaced pixel. The pixel shader compiles 2 and 18 instructions for pass one and pass two, respectively, for Shader Model 2.0. The HLSL code is shown as below. In the first pass (PS0), the only pixel shader instruction is to fetch the sampling intensity from the volume texture. In the second pass (PS1), the intensities of inner point and outer point are fetched respectively and then the displaced pixel shading is performed. Evaluated with the FX Composer from NVIDIA using the popular GeForce 6800 GT video card and v77.72 driver, the first pass uses one internal register and is executed in one instruction cycle. The approximate pixel throughput is 5.60 GigaPixels (GP)/s. In the second pass, two internal registers are used, and the pass is executed in seven instruction cycles, with an approximate pixel throughput of 0.8 GP/s. In general, the second pass takes about seven execution times as the first pass.

```
PS_OUTPUT PS0( PS_INPUT IN )
{
    PS_OUTPUT      OUT;
    OUT.color = tex3D( VolumeSampler, IN.texturecoord );
    return OUT;
}
PS_OUTPUT PS1( PS_INPUT IN )
{
    PS_OUTPUT      OUT;
    float4      isopoint, innerpoint, outerpoint;
    float4      coord;
    innerpoint = tex3D( VolumeSampler, IN.texturecoord );
    coord = IN.texturecoord - c3 * c0;
    outerpoint = tex3D( VolumeSampler, coord );
    isopoint.a = (innerpoint.a - c3.a) / (innerpoint.a - outerpoint.a);
    clamp( isopoint.a, 0.001, 1.0);
    coord = lerp( IN.texturecoord, coord, isopoint.a);
    isopoint = tex3D( VolumeSampler, coord );
    OUT.color.a = 1.0;
    isopoint.rgb = 2 * (isopoint.rgb - 0.5);
    OUT.color.rgb = saturate(dot(isopoint.rgb, c0)) * c2 + c1;
    return OUT;
}
```

3. RESULTS

We tested the algorithm on a general PC platform. The CPU had a Pentium 4 2.66G processor, 1G memory, and an ATI 9500 Pro GPU card. A 128x128x94 skull and a 512x512x193 phantom model were rendered for comparison. The results are shown in Fig. 5, Fig. 6, and Table 1. In Figs. 6 and 7, the first images for each data set are generated by the original slice-based texture volume rendering with a few number of slices. The artifacts are easy to find. The dark points result from the low sampling rate. In the second images for each dataset, a large number of slices from the slice-based rendering algorithm were used. The image quality is better than the first set, but the performance was worse due to the large amount of pixel shader operations. The third images are from our displaced pixel shading algorithm with the same number of slices as the first image. The image quality is similar with the over-sampled image, with an interactive rendering speed.

From the experiments, our proposed displaced pixel shading algorithm can remove many of the artifacts caused by the low sampling rate with the same slice number, and the resulting image quality is similar to the high sampling rate rendering. The performance of optimized displaced rendering, as indicated in Table 1, is almost the same as the original low sampling rendering and is several time faster than the high sampling rendering depends on the number of slices used. Thus, we conclude that our algorithm achieves a high quality volume rendering result with fewer slices than the traditional slice-based GPU rendering algorithm.
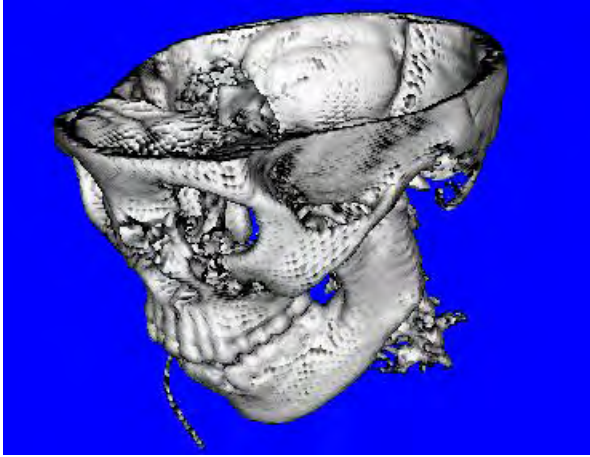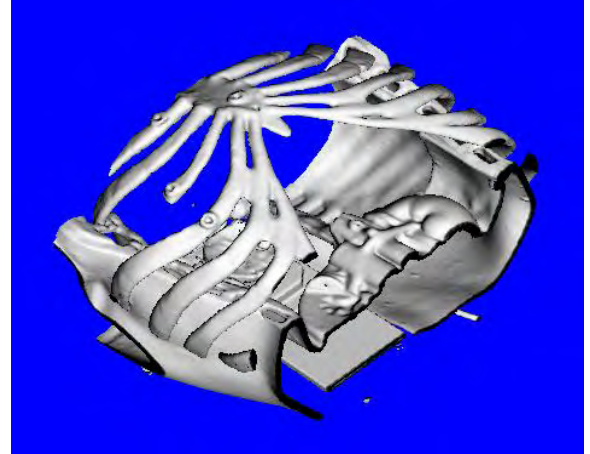

(a) Non-displaced (250 slices)
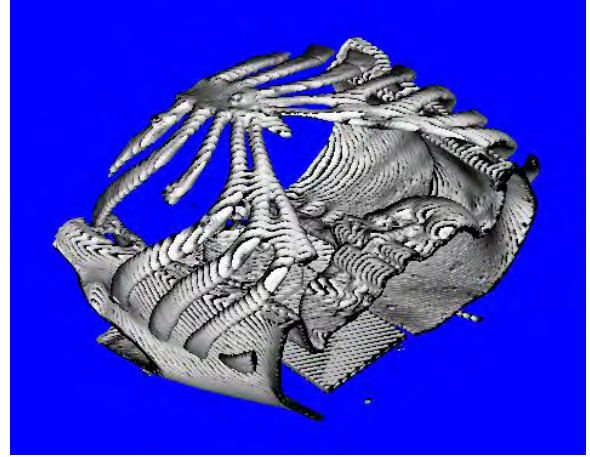
(a) Non-displaced (200 slices)
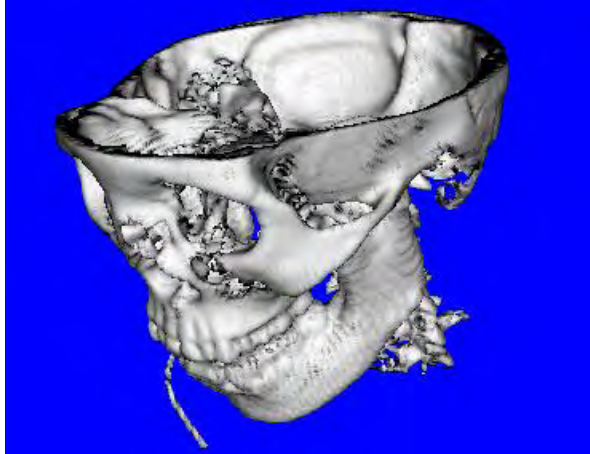
(b) Non-displaced (1000 slices)
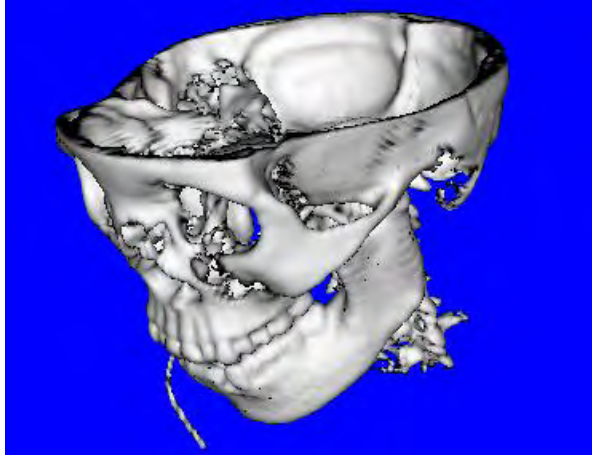

(b) Non-displaced (500 slices)
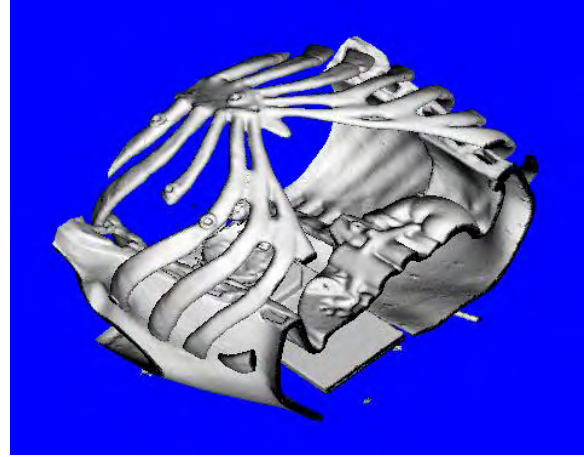
Fig. 6 (c) Displaced pixel shading (250 slices) Fig. 7 (c) Displaced pixel shading (200 slices)

Dataset		Non-displaced	Non-displaced	Displaced
128x128x94	Slice	250	1000	250
8 Bits	FPS	48	6	46
512x512x193	Slice	200	500	200
16 Bits	FPS	39	13	38

Table 1 Comparison of rendering performance

4. CONCLUSIONS

In this paper, a new GPU-based displaced pixel shading algorithm is presented to improve the image quality of the slice-based GPU-based ray casting algorithm and uses a small number of slices. Our algorithm uses several techniques – such as displaced pixel shading, multi-pass and early Z culling, and a moment-based edge detector – to render the photorealistic image in an interactive speed. The results indicate that those techniques are efficient in accelerating the rendering speed with good image quality. The future work will focus on the exploring the new released features in the DirectX 10 pipeline and produce composite and more realistic images.

ACKNOWLEDGEMENTS

REFERENCES

1. Van Gelder, A., Kwansik, K., Direct Volume Rendering with Shading via Three-Dimensional Texture, ACM symposium on Volume Visualization 96, R. Crawfis and C. Hansen, Eds., 23-30, 1996
2. Westermann, R., Ertl, T. Efficiently using graphics hardware in volume rendering applications, Computer Graphics (SIGGRAPH 98 Proceedings), 291-294, 1998
3. Meissner, M., Hoffmann, U.,, Strasser, W., Enabling classification and shading for 3D texture mapping based volume rendering using OpenGL and extensions. IEEE Visualization '99, 110-119, 1999

4. Rezk-Salama, C., Engel, K., Bauer, M., Greiner, G., Ertl, T., Interactive volume rendering on standard PC graphics hardware using multi-textures and multi-stage rasterization. SIGGRAPH / Eurographics Workshop on Graphics Hardware, 109-119, 2000

5. Engel, L., Kraus, M., Ertl, Tl. High-quality pre-integrated volume rendering using hardware-accelerated pixel shading. SIGGRAPH / Eurographics Workshop on Graphics Hardware. 2001

6. Guthe, S., Roettger, S., Schieber, A., Strasser, W., Ertl, T., High-quality unstructed volume rendering on the PC platform. ACM SIGGRAPH / Eurographics Hardware Workshop, 2002

7. Kniss, J., Premoze, S., Hansen, C., Ebert, D., Interactive translucent volume rendering and procedural modeling. Proceedings of IEEE Visualization 2002, 168-176, 2002

8. Jens Kruger, Ruduger Westermann, Acceleration Techniques for GPU-based Volume Rendering, Proceeding of IEEE Visualization 2003, 287-292, 2003

9. Simon, Green, Volume Rendering for Games, Game Development Conference 2005, 2005

10. Jason L. Mitchell, Pedro V. Sander, Applications of Explicit Early-Z Culling, SIGGRAPH 2004 – Real-time Shading Course, 2004

11. Wolfgang Engel, Shader Programming - Part III: Fundamentals of Pixel Shaders, http://www.gamedev.net/columns/hardcore/dxshader3/

12. Hui Zhang, Huazhong Shu, Limin Luo, Jean-Louis Dillenseger, A Legendre Orthogonal Moment based 3D Edge Operator, Science in China (Series G), 48(1):1-13, 2005

8.11 Zhang 2007: Treatment Planning and Image …

Reprint begins on the next page and is ten pages.

Treatment Planning and Image Guidance
for Radiofrequency Ablation of Liver Tumors

Hui Zhang[a,b], Filip Banovac[b], Stella Munuo[c],
Enrique Campos-Nanez[c], Hernan Abeledo[c], Kevin Cleary*[b]
[a]Accuray Inc, 1310 Chesapeake Ter, Sunnyvale, CA USA 94089;
[b]Imaging Science and Information Systems (ISIS) Center, Department of Radiology,
Georgetown University, 2115 Wisconsin Ave, Suite 603, Washington, DC USA, 20007;
[c]Department of Engineering Management and Systems Engineering, George Washington University,
2121 I St., Washington, DC USA 20052

ABSTRACT

Radiofrequency ablation is becoming an increasingly attractive option for minimally invasive treatment of liver tumors. In this procedure, the tumor and its margin are ablated using radiofrequency ablation probes that cover a region from 2cm to 7cm in diameter. For a large or irregularly shaped tumor, multiple ablations with overlapping probe placements are required. In this paper, we propose a treatment planning system to optimize these placements. A general optimization framework based on inverse planning methods is designed to generate the treatment plan. An objective function is defined to describe the coverage of the ablation volumes. Powell's method and simulated annealing algorithms are used to find the solution. Pre-computed mask volumes and an initial placement based on a Euclidean Distance Transform are used to speed up the computation, which can generally take a few seconds to several minutes. To ensure accurate placement of the ablation probe, we also propose a system architecture for integrating the treatment planning system with our previously developed image-guided surgery system, which uses an electromagnetic tracking device. We present some preliminary results from synthetic data to validate our treatment planning algorithm and system concept.

Keywords: Radiofrequency ablation, treatment planning, image-guided surgery system, electromagnetic tracking

1. INTRODUCTION

Radiofrequency ablation is becoming an increasingly attractive option for treatment of liver tumors. During this procedure, an ablation probe is delivered into the tumor volume; then, the probe radiates radiofrequency energy to the surrounding tissues. This energy heats the tumor and causes cell death, as shown as Figure 1. The whole procedure is minimally invasive, safe, and approved by the Food and Drug Administration (FDA) for soft-tissue ablations [1].

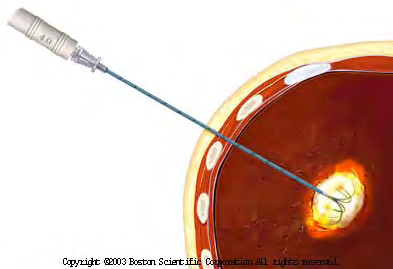Figure 1. Radiofrequency ablation of liver tumor. Image courtesy of Boston Scientific.

Although radiofrequency ablation has gained wide acceptance as a minimally invasive means to avoid laparotomy, it often has a higher recurrence rate than hepatic resection [2]. To prevent local recurrence, a margin of 0.5~1 cm beyond

---

* Email: cleary@isis.georgetown.edu, Tel 202-687-2092, Fax 202-784-3479

the tumor must be ablated; however, this treatment can be difficult in the case of a large tumor, for which the ablation probe must be placed in multiple positions to cover the whole volume. Two issues arise when trying to treat larger tumors by this method: first, it is difficult to ablate a large and irregular tumor using multiple ablations without a treatment plan; and second, image guidance methods are required to accurately place the needle and to update the treatment plan according to the actual probe placement. Therefore, two capabilities are necessary for accurate ablation: pre-operative treatment planning and intra-operative image guidance.

Several research groups have proposed different radiofrequency ablation treatment plans that address the probe placement problem. Generally there are two major approaches. The first approach is to build some pre-defined geometry models to cover the tumor, based on the assumption that the tumor's shape can be approximated by a sphere or other simple geometrical objects. The placement of those ablation spheres is then mapped to the tumor volume. Dodd and Chen created several pre-defined geometric models for tumors of various sizes [3, 4]. The second approach is to use some mathematical solutions to generate the treatment plan. Some researchers have focused on the solution to the volume coverage problem, and others have focused on precisely simulating the heat-sink effect caused by blood flow. For the volume coverage problem, an objective function is constructed to describe the tumor coverage and optimization techniques are used to find the solution, which includes the number and the positions of each ablation probe. This kind of work has focused on generating feasible treatment plans relatively quickly. Butz proposed an interesting method to find the probe placements [5]. His method uses the 3D Slicer visualization software to simulate the procedure, but the surrounding and critical tissues are not considered. Villard proposed a framework to simulate and plan the virtual radiofrequency ablation of liver tumors [6-8]. The minimal spheroid covering volume, which includes the tumor and its margin volume, is optimized. The comparison of different optimizers, such as downhill simplex and Powell's method, is also provided. Finite Element Methods (FEM) and Boundary Element Methods (BEM) are mostly used for predicting the realistic thermal distribution [2, 9-12]. These methods provide a more accurate solution, but may not be clinically useful as the computation can take a long time.

Image guidance is the second issue for the computer aided radiofrequency ablation procedure. The pre-operative treatment plan determines the image guidance during the treatment. Image guidance and validation techniques are required to guide and evaluate the clinical procedure. Nicolau used a 3D/2D registration method to guide the ablation [13]. The radio-opaque markers are extracted and the needles are tracked in real-time. Bricault introduced a Computer Aided Diagnosis (CAD) tool to analyze the treated tumor, as a means of early detection for local recurrences [14].

In this paper, we propose a system for radiofrequency ablation procedures that both plans the placement of probes for treatment and incorporates image guidance using an electromagnetic tracking device. The treatment planning system is based on the inverse planning method. An objective function that considers several factors, such as overlapped volume and an Organ At Risk (OAR) volume, is constructed. Powell's method and simulated annealing algorithms are used to find the solution. Some acceleration techniques, such as pre-computed ablation volume and initial placement approximation based on the Euclidean Distance Transform (EDT), are used to improve the algorithm performance, a critical step for a practical image guidance system. The treatment plan can also be validated and updated during the procedure. The image guidance system uses an electromagnetic tracing device to validate and update the position of the probes during the ablation procedure. The goal of our proposed system is to increase the precision and success rate of clinical procedures.

2. METHOD

2.1 Treatment Planning

A successful radiofrequency ablation treatment plan should:

- Minimize the number of ablations and probe trajectories to reduce treatment cost and patient discomfort;
- Minimize the treatment volume, which includes the original tumor and its margin volume;
- Protect the critical organs around the tumor;
- Perform well enough to be able to update the plan during the treatment procedure.

The two major techniques for generating radiotherapy treatment plans are forward planning and inverse planning. In forward planning, the treatment plan is generated by the physicians, and the solution is simulated and displayed on the computer. The dose distribution information, such as a Dose Volume Histogram (DVH) and the visualization results, is

used to make the decision by the physicians. The physicians can change the treatment plan until an acceptable solution is achieved. This type of forward planning is like manual planning with computer-assisted feedback. In inverse planning, which we propose to use here, the desired dose distribution is specified by the physicians. An objective function is built by analyzing the treatment requirements. The final solution can then be computed by different optimization techniques. Compared with forward planning, the performance and the precision of the treatment plan are all improved.

Although these two planning methods were originally developed for radiotherapy, they can also be used for radiofrequency ablation. In the inverse planning method proposed here, the segmented tumor and the margin, which should be fully ablated, are interpreted as the 100 percent dose contour. The minimal coverage volume and the number of ablations are then optimized.
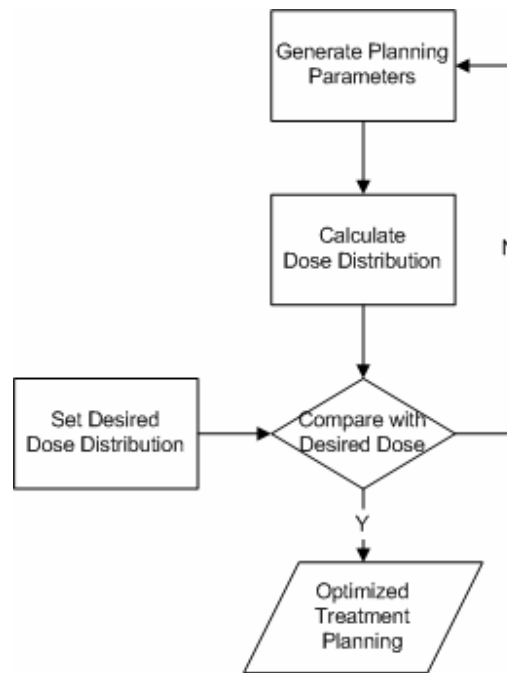

Figure 2. Diagram of inverse planning.

An objective function is required to define the treatment problem. We divide the volume in the target image into several categories:

- Target-ablation volume V_t: the segmented tumor and its margin. This volume should be fully covered in the treatment plan to prevent local recurrence of the tumor.

- Normal-ablated volume V_a: the part inside the target ablation volume covered by one ablation only. This volume should be maximized for an optimum treatment plan.

- Multiple-ablated volume V_m: the part inside the target ablation volume covered by multiple ablations. Although there is no harm in ablating a portion of the tumor volume multiple times, it is better to minimize this region, to keep the treatment plan efficient.

- Effective-ablated volume V_e: the volume inside the target ablation volume that is ablated. This volume is a sum of normal-ablated volume and multiple-ablated volume. $V_e = V_a + V_m$.

- Non-ablated volume V_n: the part inside the target ablation volume that is not covered by any ablation. This volume must be minimized to guarantee the death of all tumor cells. $V_n = V_t - V_e$.
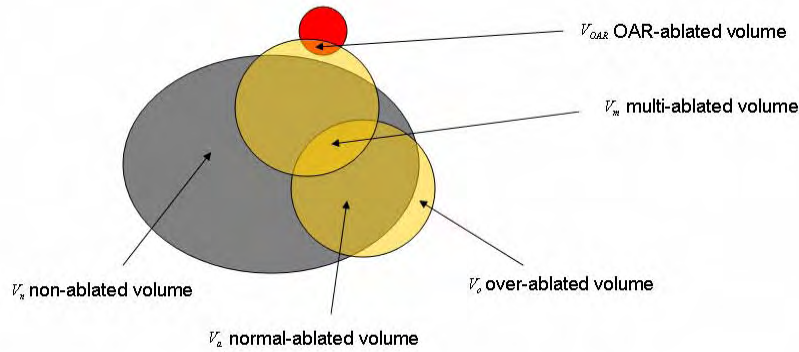
- Over-ablated volume V_o: the part outside the target ablation volume that is covered by the ablations. This volume measures how much healthy tissue is ablated by the treatment. This volume must also be minimized to reduce damage to surrounding organs. The impact of this value on the plan differs by ablation procedure. For liver ablation, this volume can be a bit large, but for cardiac ablation, it must be small.

- OAR-ablated volume V_{OAR}: the part inside the Organ At Risk volume that is ablated. This volume should be considered during the ablation and must be kept to zero to protect the patient.

All these volumes appear in Figure 3. V_t is fixed and determined by the tumor segmentation and the computation of its margin volume. V_a and V_e should be maximized. V_m, V_n, V_o, and V_{OAR} should each be minimized and have different impact weights. Based on this classification, a generic object function is defined as:

$$O(p) = \frac{1}{V_t}\left(w_a \cdot V_a + w_m \cdot V_m + w_n \cdot V_n + w_o \cdot V_o + w_{OAR} \cdot V_{OAR}\right)$$

$$V_t = V_a + V_m + V_n$$

where $p = \left(x_1, y_1, z_1, r_1, \cdots, x_N, y_N, z_N, r_N\right)$. N is the number of ablations. p is the parameter vector that contains $4 \times N$ unknowns. Among those unknowns, x, y, z is the center of the ablation sphere in the continuous space and r is the radius in the discrete spacing defined by the specifications of ablation probes from the manufacturers. r can be extended to three dimensions to simulate an ellipsoid ablation probe. w_a, w_m, w_n, w_o, and w_{OAR} are the weights of the different volume types and reflect the different impacts.

Figure 3. Design of objection function for radiofrequency ablation

This objective function is a configurable measurement. The weights are adjustable for different purposes and measurements. Butz and Villard's measurements can also be simulated through this objective function [5-8]. If we set $w_a = w_m$, $w_n = 0$, and $w_{OAR} = 0$, the formula is similar to Butz's mathematical expression for optimization. If we set $w_a = w_m = 0$, $w_n \to \infty$, and $w_{OAR} = 0$, the effect is close to Villard's optimization work. In this case, $w_n \to \infty$

means 100% coverage of the target volume. A weight of w_o is used to minimize the ablation coverage. Different weights will affect the behavior and final solution of the inverse planning problem, so they should be set carefully. Some general rules are:

- Set w_n to a high value to guarantee full coverage of the target and successful treatment.

- Set w_{OAR} to an absolute high value to keep the procedure safe and to protect the patient.

- Increase w_o to reduce the damage to surrounding organs. This value changes according to different operation positions and procedures, such as liver ablation and cardiac ablation.

- Increase w_m to make the procedure more effective.

Our treatment planning system uses Powell's method and simulated annealing algorithms to compute a solution. To generate a feasible treatment plan, extensive iterations may be required to locate the minimal value of the objective function. In a single iteration, the target volume and the ablation spheres defined by the parameters should be masked to measure those different volumes. The computation time in a single iteration should be reduced to improve overall performance. We propose two methods to accelerate the treatment plan generation: pre-computed mask volume and initial placement approximation by Euclidean Distance Transform.

In a single iteration, the voxel around the target volume should be measured to determine whether it is ablated or not. In the paper by Butz et al. [5], this measurement was computed by comparing the distance from each voxel to the ablation center with the radius of the ablation sphere. However, this kind of computation must be performed each time and thus is computationally expensive. To reduce the computational time, pre-computed mask volumes are generated for different sizes of ablation probes. The pre-computed value of '1' or '0' indicates whether this voxel is inside or outside of the ablation sphere. The pre-computed volumes of the different-sized ablation probes are used to mask the voxels in the target volume. The computationally expensive measurement is replaced by a single comparison operation. This method is very useful in accelerating the volume measurement.

In our framework, the optimizer is trying to locate the minimal value of the objective function in the parameter space. The start position in the parameter space determines the solution that is found by the algorithm. In particular, this can affect solution quality since the objective function can have multiple local optima. The initial position affects also the optimization speed and iteration time. If the initial position is close to the final solution, the optimization procedure will be greatly accelerated. We propose an initial placement approximation based on the Euclidean Distance Transform. A Euclidean distance map is generated for the target volume to indicate the distance from the internal voxel to the contour. Saito's algorithm is used to generate the Euclidean distance map [15]. Based on the distance information, the initial start position is located using the algorithm, as explained in the box below.

A) Generate the Euclidean distance map for the voxels inside the target volume. $D(i, j, k)$ indicates the distance intensity of the voxel located at $p(i, j, k)$;

B) Compute the geometric center C_{target} of the target volume;

C) According to the radius of the specific ablation probe, generate the internal possible contour set using the voxels with distance intensity greater than that radius. $\Phi = \{p(i, j, k)\}$ where $D(i, j, k) > r$;

D) For all the voxels inside the contour, find the voxel that has the largest distance to C_{target};

E) Use this voxel as the ablation probe candidate and update the possible contour set to remove all the voxels covered by the sphere centered at this candidate point;

F) Go back to step C) until all the voxels in the contour are removed.

Candidate points are generated from this approximation algorithm and used as the start positions. In some cases, those initial positions are very close to the final solution and the whole optimization is accelerated. The whole procedure is

shown in Figure 4. Figure 4(a) shows the initial tumor contour, the internal candidate contour based on distance transform, and also the geometry center. Figure 4(b) shows how the first candidate point is located, and Figure 4(c) shows how the other candidates are located consequently.
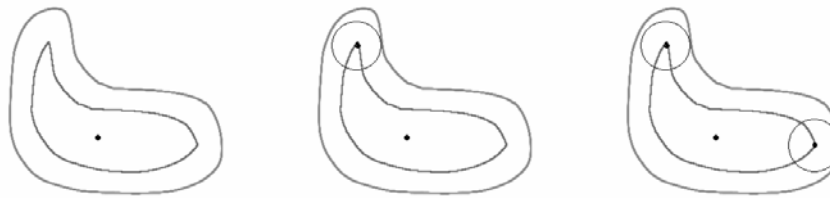
(a) Define initial tumor contour, internal candidate contour, and geometric center
(b) Find the first voxel in the internal candidate contour
(c) Continue to find other candidates
Figure 4. Initial placement approximation using Euclidean Distance Transform.

2.2 Image Guidance using Electromagnetic Tracking Device

Image-guided systems have been commercially available for over ten years now, but most of these systems are based on optical tracking of surgical instruments. Optical tracking provides high precision and a large field of view, but its application is restricted by the line-of-sight requirement and inability to track the tip of flexible and bending surgical tools such as biopsy needles and guidewires. Electromagnetic tracking devices are capable of tracking internal organs and invasive surgical tools, but their precision and performance is not as good as those of optical tracking devices. The distortion of the electromagnetic field caused by metal objects may also degrade the accuracy of electromagnetic tracking devices.

Our research group at Georgetown University has been integrating electromagnetic tracking devices with an image-guided system as shown in Figure 5 and has used this system in several studies, such as for liver biopsy [16]. Different algorithms were investigated to track the motion of an internal organ based on single or multiple internal trackers [17]. In the image-guided system, the position of the needle tip is overlaid with the pre-operative planning volume to track the position of the instrument in real-time. Based on this information, the radiofrequency ablation needle can be guided by the system to reach the planned positions and the thermal energy can be delivered to destroy the tumor. The actual ablation positions during the procedure can be recorded to validate the treatment, or potentially update the treatment plan after each ablation to account for any errors in probe placement.
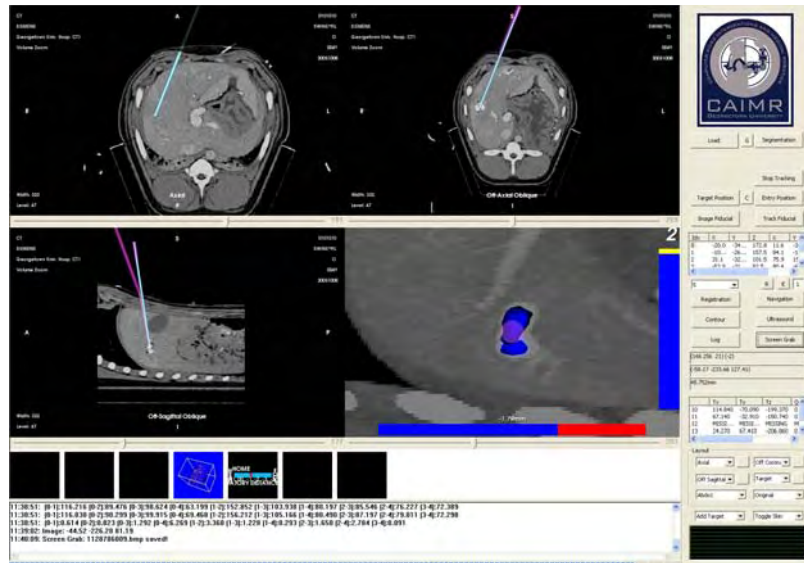

Figure 5. Image-guided surgery system using electromagnetic tracking device.
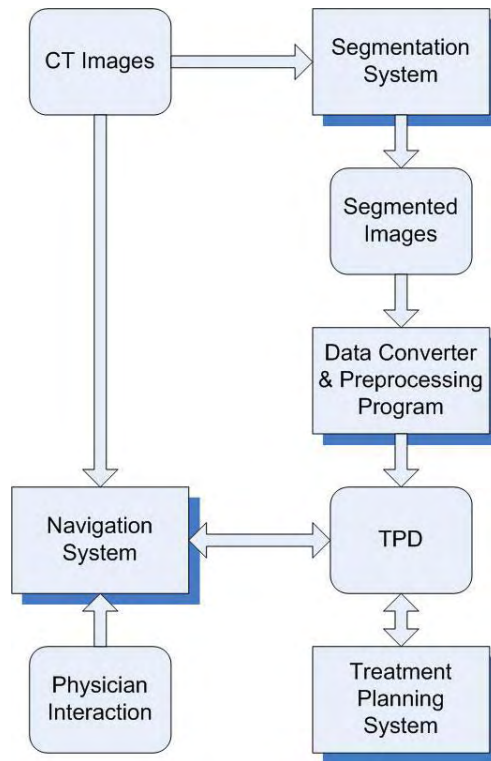
Figure 6. System workflow for our image-guided radiofrequency ablation system. TPD means Treatment Planning Data.

2.3 System Architecture and Workflow

The architecture of the proposed system is shown in Figure 6. The pre-operative CT images are input into the segmentation system and segmented by the physicians. The margin volume of the segmented image is then calculated using the pre-processing program. The processed image is then converted to the treatment data format, which uses different values to identify the internal tumor voxels, boundary tumor voxels, margin voxels, and critical voxels. A data format description is shown as Table 1. The treatment planning system reads the treatment data and generates the treatment plan, which includes the number, radiuses, and positions of the ablation probes. The treatment plan is input into our image guidance system for the procedure guidance and tracking. The updated information from the clinical ablation is exchanged between the treatment planning system and image guidance system for real-time update and validation. With this workflow, the segmentation system, treatment planning system, and image-guidance system are integrated together. The workflow gives each sub-system maximum flexibility.

```
! Part 1: Scalar information
radius_1,radius_2                 ! Ablation radius
grid_size                         ! Grid size
angular_res                       ! Angular resolution
max_angle                         ! Maximum angle between directions
max_ablation              ! Maximum no. of ablations
max_rays                  ! Maximum no. of directions
max_punctures             ! Maximum no. of punctures
! Part 2: Voxel information
x1,y1,z1,VOXEL_FLAG
x2,y2,x2,VOXEL_FLAG
...
```

Table 1. Treatment planning data format.

3. RESULTS

Some synthetic data were generated for an initial algorithm validation and system test. Chen proposed some geometric models for tumors whose sizes range from 4.0-4.3cm, 4.4-4.6cm, 4.7-5.1cm, 5.2-5.4cm, 5.5-5.6cm and 5.7-6.5cm [4]. A 0.5cm margin is applied to those tumors and the number of required ablation spheres is provided. Following his example, we generated some spherical tumor data with radiuses of 4.0cm, 4.3cm, 4.7cm, 5.1cm, 5.5cm, 5.6cm, 5.7cm and 6.5cm, each with 0.5cm margin. Those data were input into our treatment planning system to generate the plan and the results were compared with the pre-defined geometry models. To simulate the segmented medical image data, we located those tumor targets in a 256x256x256 volume with spacing of 1mm in all directions. 5.0cm ablation probes were used for the plan. The generated plan results are shown in Table 2 on the next page.

Since the tumor volume is discretized into voxels, the computed target coverage is approximated by those voxels with an accuracy of 1mm^3. Only the treatment plans whose approximate target coverage was greater than 95% were listed. For large tumors, the number of ablations required to fully cover the targets was lower in our treatment planning system than in the geometry models. In particular, for the tumor with the diameter of 5.7cm, our treatment plan required only 8 ablations, while 12 ablations were required in the geometry model. The computation time varied from 0.344 seconds to 53.844 seconds in those cases – fast enough for the pre-operative treatment planning and also for the intra-operative plan update. The hardware configuration was a Pentium 4 3.4GHz processor with 1GB RAM. Powell's algorithm was used for the optimization. Simulated annealing algorithm was also implemented, but the results did not show much improvement in those cases while the algorithm required much more computational time. Figure 7 shows two example cases: a 5.7cm tumor with a 0.5cm margin covered by eight 5cm ablations, and a 6.5cm tumor with a 0.5cm margin covered by ten 5cm ablations.
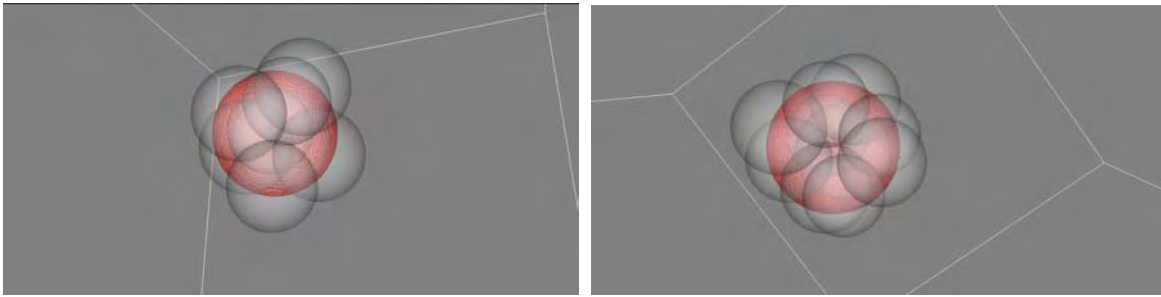

Figure 7. Transparent view of sphere overlapping.
(a) 5.7cm tumor with 8 ablations.   (b) 6.5cm tumor with 10 ablations.

To verify the plan's accuracy, three synthetic tumor data sets were constructed. Those data contained one, two, and three spherical tumors, each with a diameter of 4.0cm. The known solutions for the treatment plan were the centers of those spheres. The results generated from our treatment planning system are shown as Table 3 on the next page. The general errors of the probe positions were less than 1.9mm for those large irregular tumors. This result demonstrates that our treatment plans are feasible for the system.

4. CONCLUSIONS

Our paper proposes a radiofrequency ablation treatment planning system that places ablation probes to cover the tumor. A generic optimization framework with several unique acceleration techniques is applied to compute the treatment plan. We also propose an architecture and workflow for a complete image-guided treatment planning system that uses electromagnetic tracking to help carry out the treatment plan.

In future work, we need to test our system and algorithm on clinical tumor data. In addition to optimizing the placement of the ablation probes, the probe trajectory should also be optimized to minimize damage to healthy tissue. The weighting parameters necessary to optimize the results of specific procedures should also be identified.

Tumor Diameter (cm)	Target Diameter (cm)	Number of Ablations. (Geometric Model)	Number of Ablations (Treatment Planning)	Approximate Target Coverage (%)	Number of Iterations	Computation Time (s)
4.0	5.0	1	1	100.000	120	0.344
4.3	5.3	4	2	98.662	252	1.062
			3	99.890	378	1.812
			4	100.000	756	4.032
4.7	5.7	6	3	97.813	720	3.718
			4	99.960	960	5.782
			5	100.000	900	6.297
5.1	6.1	6	4	99.115	1212	7.656
			5	99.874	2415	17.547
			6	100.000	1818	15.516
5.5	6.5	8	4	96.752	1236	8.422
			5	98.648	1860	14.093
			6	99.990	3347	28.531
			7	100.000	2164	19.844
5.6	6.6	8	4	95.889	1237	8.516
			5	98.269	1860	14.062
			6	99.975	2610	22.406
			7	100.000	2163	20.875
5.7	6.7	12	5	97.737	3090	23.375
			6	99.920	2574	22.124
			7	99.997	2584	24.655
			8	100.000	1948	20.361
6.5	7.5	12	6	97.614	2615	23.890
			7	99.064	3470	35.750
			8	99.732	4452	48.391
			9	99.984	4462	53.844
			10	100.000	3116	40.531

Table 2. Treatment planning results for spherical tumors.

Number. of Tumors	Number of Iterations	Computation Time (s)	Center Positions (mm)	Plan Positions (mm)	Error (mm)
1	120	0.344	(110.0, 110.0, 110.0)	(110.0, 110.0, 110.0)	0.0
2	240	0.968	(110.0, 110.0, 110.0)	(110.8, 110.0, 110.0)	0.2
			(100.0, 110.0, 110.0)	(99.2, 110.0, 110.0)	0.8
3	913	4.656	(110.0, 110.0, 110.0)	(110.2, 111.3, 111.2)	1.8
			(90.0, 110.0, 110.0)	(89.8, 111.3, 110.0)	1.3
			(110.0, 90.0, 110.0)	(110.2, 91.5, 111.2)	1.9

Table 3. Accuracy comparison of treatment plans.

ACKNOWLEDGMENTS

REFERENCES

1. SA Curly, F Izzo, P Delrio, et al. "Radiofrequency ablation of unresectable primary and metastatic hepatic malignancies: result in 123 patients", *Ann Surg* 230:1-8, 1999
2. TWH. Sheu, CW Chou, SF Tsai and PC Liang, "Three-dimensional analysis for radio-frequency ablation of liver tumor with blood perfusion effect", *Computer Methods in Biomechanics and Biomedical Engineering*, 8(4):229-240, 2005
3. GD Dodd, MS Frank, M Aribandi, S Chopra, KN Chintapalli, "Radiofrequency thermal ablation: computer analysis of the size of the thermal injury created by overlapping ablations", *American Journal of Roentgenol* 177:777-782, 2001
4. MH Chen, W Yang, K Yan, MW Zou, L Solbiati, JB Liu, Y Dai, "Large liver tumors: protocol for radiofrequency ablation and its clinical application in 110 patients: mathematic model, overlapping mode, and electrode placement process", *Radiology* 232:260-271, 2004
5. T Butz, SK Warfield, K Tuncali, SG Silverman, E Sonnenberg, FA Jolesz, R Kikinis, "Pre- and intra- operative planning and simulation of percutaneous tumor ablation", *MICCAI 2000, Lecture Notes In Computer Science*, Vol. 1935, 317-326, 2000
6. C Villard, L Soler, A Gangi, "Radiofrequency ablation of hepatic tumors: simulation, planning, and contribution of virtual reality and haptics", *Computer Methods in Biomechanics and Biomedical Engineering*, 8(4):215-227, 2005
7. C Villard, L Soler, A Gangi, D Mutter, J Marescaux, "Towards realistic radiofrequency ablation of hepatic tumors 3D simulation and planning" *SPIE Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display*, edited by Galloway, Robert L., Jr. Proceedings of the SPIE, Vol. 5367, 586-595, 2004
8. C Villard, L Soler, N Papier, V Angus, S Thery, A Gangi, D Mutter, J Marescaux, "Virtual Radiofrequency Ablation of Liver Tumors", *IS4TM 2003, Lecture Notes In Computer Science* Vol. 2673, 366-374, 2003
9. MK Jain, PD Wolf, "A Three-Dimenional Finite Element Modal of Radiofrequency Ablation with Blood Flow and its Experimental Validation", *Annals of Biomedical Engineering*, 28:1075-1084, 2000
10. J Gopalakrishnan, "A Mathematical Modal for Irrigated Epicardial Radiofrequency Ablation", *Annals of Biomedical Engineering*, 30:884-893, 2002
11. S Tungjitkusolmun, ST Staelin, D Haemmerich, JZ Tsai, H Cao, JG Webster, FT Lee, Jr., DM. Mahvi, VR Vorperian, "Three-Dimensional Finite-Element Analyses for Radio-Frequency Hepatic Tumor Ablation", *IEEE Transactions on Biomedical Engineering*, 49(1):3-9, 2002
12. CCR Chen, MI Miga, RL Galloway, "Optimizing Needle Placement in Treatment Planning of Radiofrequency Ablation", *Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display*, edited by Cleary, Kevin R.; Galloway, Robert L., Jr. Proceedings of the SPIE, Vol. 6141, 632-638, 2006
13. S Nicolau, A Garcia, X Pennec, L Soler, N Ayache, "An augmented reality system to guide radio-frequency tumor ablation", *Computer Animation and Virtual Worlds*, 16(1):1-10, 2005
14. I Bricault, R Kikinis, EV Sonnenberg, K Tuncali, SG Silverman, "3D Analysis of Radiofrequency-Ablated Tumors in Liver: A Computer-Aided Diagnosis Tool for Early Detection of Local Recurrences", *MICCAI 2004*, *Lecture Notes In Computer Science* Vol. 3217, 1042-1043, 2004
15. T Saito, J Toriwaki, "Algorithms of three dimensional Euclidean distance transformation and extended digital Voronoi diagram, and analysis of human liver section images", *The Journal of the Institute of Image Electronics Engineers of Japan*, 21(5):468-474, 1992
16. F Banovac, E Wilson, H Zhang, K Cleary, "Needle biopsy of anatomically unfavorable liver lesions with an electromagnetic navigation assist device in a computed tomography environment", *J Vasc Interv Radiol*, 17(10):1671-1675, 2006
17. H Zhang, F Banovac, R Lin, N Glossop, BJ Wood, D Lindisch, E Levy, K Cleary, "Electromagnetic tracking for abdominal interventions in computer aided surgery", *Comput Aided Surg,* 11(3):127-36, 2006